
JMIR Formative Research

Impact Factor (2022): 2.2

Volume 5 (2021), Issue 11 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

- A Transdiagnostic Self-management Web-Based App for Sleep Disturbance in Adolescents and Young Adults: Feasibility and Acceptability Study ([e25392](#))
Nicole Carmona, Aleksandra Usyatynsky, Samlau Kutana, Penny Corkum, Joanna Henderson, Kelly McShane, Colin Shapiro, Souraya Sidani, Jennifer Stinson, Colleen Carney. 7
- Organizational Readiness for Implementing an Internet-Based Cognitive Behavioral Therapy Intervention for Depression Across Community Mental Health Services in Albania and Kosovo: Directed Qualitative Content Analysis ([e29280](#))
Asmae Doukani, Arlinda Cerga Pashoja, Naim Fanaj, Gentiana Qirjako, Andia Meksi, Sevim Mustafa, Christiaan Vis, Juliane Hug. 25
- Acceptability, Engagement, and Exploratory Outcomes of an Emotional Well-being App: Mixed Methods Preliminary Evaluation and Descriptive Analysis ([e31064](#))
Amelia Eisenstadt, Shaun Liverpool, Athina-Marina Metaxa, Roberta Ciuvat, Courtney Carlsson. 39
- Understanding the Social Determinants of Mental Health of Undergraduate Students in Bangladesh: Interview Study ([e27114](#))
Ananya Bhattacharjee, S Haque, Md Hady, S Alam, Mashfiqui Rabbi, Muhammad Kabir, Syed Ahmed. 55
- A Mobile App to Enhance Behavioral Activation Treatment for Substance Use Disorder: App Design, Use, and Integration Into Treatment in the Context of a Randomized Controlled Trial ([e25749](#))
Catherine Paquette, Dillon Rubalcava, Yun Chen, Deepika Anand, Stacey Daughters. 71
- A Video Self-Modeling Intervention Using Virtual Reality Plus Physical Practice for Freezing of Gait in Parkinson Disease: Feasibility and Acceptability Study ([e28315](#))
Lina Goh, Natalie Allen, Naseem Ahmadpour, Kaylena Ehgoetz Martens, Joeeun Song, Lindy Clemson, Simon Lewis, Hamish MacDougall, Colleen Canning. 83
- Feasibility and Acceptability of a Digital Patient-Reported Outcome Tool in Routine Outpatient Diabetes Care: Mixed Methods Formative Pilot Study ([e28329](#))
Soren Skovlund, Lise Troelsen, Lise Noergaard, Anna Pietraszek, Poul Jakobsen, Niels Ejksjaer. 100
- Usability Evaluation of an Offline Electronic Data Capture App in a Prospective Multicenter Dementia Registry (digiDEM Bayern): Mixed Method Study ([e31649](#))
Michael Reichold, Miriam Heß, Peter Kolominsky-Rabas, Elmar Gräßel, Hans-Ulrich Prokosch. 122
- Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland ([e29873](#))
Kristjana Thorarinsdottir, Emily Holmes, Johann Hardarson, Unnur Hedinsdottir, Marie Kanstrup, Laura Singh, Arna Hauksdottir, Thorhildur Halldorsdottir, Berglind Gudmundsdottir, Unnur Valdimarsdottir, Edda Thordardottir, Beau Gamble, Andri Bjornsson. 139

<p>Comparing Email, SMS, and Concurrent Mixed Modes Approaches to Capture Quality of Recovery in the Perioperative Period: Retrospective Longitudinal Cohort Study (e25209) Jamie Romeiser, James Cavalcante, Deborah Richman, Sunitha Singh, Xiaohui Liang, Allison Pei, Samanvaya Sharma, Zoe Lazarus, Tong Gan, Elliott Bennett-Guerrero.</p>	154
<p>A Virtual Community for Disability Advocacy: Development of a Searchable Artificial Intelligence–Supported Platform (e33335) Christo El Morr, Pierre Maret, Fabrice Muhlenbach, Dhayananth Dharmalingam, Rediet Tadesse, Alexandra Creighton, Bushra Kundi, Alexis Buettgen, Thumeka Mgwigwi, Serban Dinca-Panaitescu, Enakshi Dua, Rachel Gorman.</p>	168
<p>Exploring the Barriers to and Motivators for Using Digital Mental Health Interventions Among Construction Personnel in Nigeria: Qualitative Study (e18969) Janet Nwaogu, Albert Chan, John Naslund, Carol Hon, Christopher Belonwu, Jackie Yang.</p>	180
<p>A Fully Collaborative, Noteless Electronic Medical Record Designed to Minimize Information Chaos: Software Design and Feasibility Study (e23789) Jackson Steinkamp, Abhinav Sharma, Wasif Bala, Jacob Kantrowitz.</p>	193
<p>Home-Based Exercise Program for Patients With Combined Advanced Chronic Cardiac and Pulmonary Diseases: Exploratory Study (e28634) Cyrille Herkert, Lidwien Graat-Verboom, Judith Gilsing-Fernhout, Manon Schols, Hareld Kemps.</p>	209
<p>A Text Message Intervention for Adolescents With Depression and Their Parents or Caregivers to Overcome Cognitive Barriers to Mental Health Treatment Initiation: Focus Groups and Pilot Trial (e30580) Brian Suffoletto, Tina Goldstein, David Brent.</p>	223
<p>A Web-Based, Population-Based Cirrhosis Identification and Management System for Improving Cirrhosis Care: Qualitative Formative Evaluation (e27748) Sarah Javier, Justina Wu, Donna Smith, Fasiha Kanwal, Lindsey Martin, Jack Clark, Amanda Midboe.</p>	235
<p>Validation of a Mobile Health Technology Platform (FeverTracker) for Malaria Surveillance in India: Development and Usability Study (e28951) Ipsita Pal Bhowmick, Dibyajyoti Chutia, Avinash Chouhan, Nilay Nishant, P Raju, Kanwar Narain, Harpreet Kaur, Rocky Pebam, Jayanta Debnath, Rabindra Tripura, Kongkona Gogoi, Suman Ch Nag, Aatreyee Nath, Debabrata Tripathy, Jotish Debbarma, Nirapada Das, Ujjwal Sarkar, Rislyn Debbarma, Rajashree Roy, Bishal Debnath, Dipanjan Dasgupta, Suraj Debbarma, Kamal Joy Tripura, Guneram Reang, Amit Sharma, Manju Rahi, Jyoti Chhibber-Goel.</p>	247
<p>Development of a Severity Score and Comparison With Validated Measures for Depression and Anxiety: Validation Study (e30313) William Lynch, Michael Platt, Adam Pardes.</p>	259
<p>Public Perceptions of Diabetes, Healthy Living, and Conversational Agents in Singapore: Needs Assessment (e30435) Dhakshenya Dhinakaran, Thirunavukkarasu Sathish, Tobias Kowatsch, Konstadina Griva, James Best, Lorraine Tudor Car.</p>	280
<p>Seeking Mental Health Support Among College Students in Video-Based Social Media: Content and Statistical Analysis of YouTube Videos (e31944) Bogeum Choi, Heejun Kim, Jina Huh-Yoo.</p>	295
<p>Formative Study of Mobile Phone Use for Family Planning Among Young People in Sierra Leone: Global Systematic Survey (e23874) Emeka Chukwu, Sonia Gilroy, Kojo Addaquay, Nki Jones, Victor Karimu, Lalit Garg, Kim Dickson.</p>	307

Dropout From an Internet-Delivered Cognitive Behavioral Therapy Intervention for Adults With Depression and Anxiety: Qualitative Study ([e26221](#))
 Kate Lawler, Caroline Earley, Ladislav Timulak, Angel Enrique, Derek Richards. 322

The Online Patient Satisfaction Index for Patients With Low Back Pain: Development, Reliability, and Validation Study ([e21462](#))
 Tamana Afzali, Henrik Lauridsen, Janus Thomsen, Jan Hartvigsen, Martin Jensen, Allan Riis. 341

Game Plan, a Web-Based Intervention to Improve Adherence and Persistence to HIV Pre-exposure Prophylaxis and Reduce Heavy Drinking in Gay, Bisexual, and Other Men Who Have Sex With Men: Usability and User Experience Testing ([e31843](#))
 Tyler Wray, Philip Chan, Christopher Kahler. 353

The 4 Youth By Youth mHealth Photo Verification App for HIV Self-testing in Nigeria: Qualitative Analysis of User Experiences ([e25824](#))
 David Oladele, Juliet Iwelunmor, Titilola Gbajabiamila, Chisom Obiezu-Umeh, Jane Okwuzu, Ucheoma Nwaozuru, Adesola Musa, Ifeoma Idigbe, Kadija Tahlil, Weiming Tang, Donaldson Conserve, Nora Rosenberg, Agatha David, Joseph Tucker, Oliver Ezechi. 368

A Two-Minute Walking Test With a Smartphone App for Persons With Multiple Sclerosis: Validation Study ([e29128](#))
 Pim van Oirschot, Marco Heerings, Karine Wendrich, Bram den Teuling, Frank Dorssers, René van Ee, Marijn Martens, Peter Jongen. 380

Evaluating a Strengths-Based mHealth Tool (MyStrengths): Explorative Feasibility Trial ([e30572](#))
 Stian Jessen, Jelena Mirkovic, Elanor Halvorsen Brendmo, Lise Solberg Nes. 399

A Technology-Based Pregnancy Health and Wellness Intervention (Two Happy Hearts): Case Study ([e30991](#))
 Tamara Jimah, Holly Borg, Priscilla Kehoe, Pamela Pimentel, Arlene Turner, Sina Labbaf, Milad Asgari Mehrabadi, Amir Rahmani, Nikil Dutt, Yuqing Guo. 426

A Smoking Cessation App for Nondaily Smokers (Version 2 of the Smiling Instead of Smoking App): Acceptability and Feasibility Study ([e29760](#))
 Bettina Hoepfner, Kaitlyn Siegel, Hannah Carlon, Christopher Kahler, Elyse Park, Susanne Hoepfner. 442

Teleassessments for Enrollment of Adults With Physical or Mobility Disability in a Home-Based Exercise Trial in Response to COVID-19: Usability Study ([e29799](#))
 Jereme Wilroy, Byron Lai, Madison Currie, Hui-Ju Young, Mohanraj Thirumalai, Tapan Mehta, John Giannone, James Rimmer. 459

Integrating SMS Text Messages Into a Preventive Intervention for Postpartum Depression Delivered via In-Home Visitation Programs: Feasibility and Acceptability Study ([e30995](#))
 Alinne Barrera, Jaime Hamil, Darius Tandon. 467

Informing Content and Feature Design of a Parent-Focused Human Papillomavirus Vaccination Digital Behavior Change Intervention: Synchronous Text-Based Focus Group Study ([e28846](#))
 Elisabeth Becker, Ross Shegog, Lara Savas, Erica Frost, C Healy, Stanley Spinner, Sally Vernon. 479

Guidance for Implementing Video Consultations in Danish General Practice: Rapid Cycle Coproduction Study ([e27323](#))
 Amanda Sandbæk, Line Christensen, Lotte Larsen, Nina Primholdt Christensen, Frida Kofod, Ann Guassora, Camilla Merrild, Elisabeth Assing Hvidt. 495

Development of a Wearable Biocueing App (Sense-IT) Among Forensic Psychiatric Outpatients With Aggressive Behavior: Design and Evaluation Study ([e29267](#))
 Annemieke ter Harmsel, Thimo van der Pol, Lise Swinkels, Anna Goudriaan, Arne Popma, Matthijs Noordzij. 509

A Phenotyping Algorithm to Identify People With HIV in Electronic Health Record Data (HIV-Phen): Development and Evaluation Study (e28620)	
Sarah May, Thomas Giordano, Assaf Gottlieb.	524
mHealth-Based Health Promotion Intervention to Improve Use of Maternity Care Services Among Women in Rural Southwestern Uganda: Iterative Development Study (e29214)	
Esther Atukunda, Lynn Matthews, Angella Musiimenta, Godfrey Mugenyi, Samuel Mugisha, Norma Ware, Celestino Obua, Mark Siedner.	5
	3
	6
Patients' Use of Mobile Health for Self-management of Knee Osteoarthritis: Results of a 6-Week Pilot Study (e30495)	
Brittany Shewchuk, Lee Green, Tanya Barber, Jean Miller, Sylvia Teare, Denise Campbell-Scherer, Kelly Mrklas, Linda Li, Nancy Marlett, Tracy Wasylak, Elena Lopatina, Deirdre McCaughey, Deborah Marshall.	550
Contemporary English Pain Descriptors as Detected on Social Media Using Artificial Intelligence and Emotion Analytics Algorithms: Cross-sectional Study (e31366)	
Ming Tan, Charlene Goh, Hee Tan.	568
Using Patient-Generated Health Data From Twitter to Identify, Engage, and Recruit Cancer Survivors in Clinical Trials in Los Angeles County: Evaluation of a Feasibility Study (e29958)	
Katja Reuter, Praveen Angyan, NamQuyen Le, Thomas Buchanan.	582
Preferences for Using a Mobile App in Sickle Cell Disease Self-management: Descriptive Qualitative Study (e28678)	
Tilicia Mayo-Gamble, Delores Quasie-Woode, Jennifer Cunningham-Erves, Margo Rollins, David Schlundt, Kemberlee Bonnet, Velma Murry.	5
	9
	7
Body Image Distress and Its Associations From an International Sample of Men and Women Across the Adult Life Span: Web-Based Survey Study (e25329)	
Alyssa Milton, Ashlea Hambleton, Anna Roberts, Tracey Davenport, Anna Flego, Jane Burns, Ian Hickie.	612
Role of 18F-Fluorodeoxyglucose–Positron Emission Tomography/Computed Tomography Imaging in the Prediction of Prognosis in Patients With Indolent Lymphoma: Prospective Study (e24936)	
Nawal AlShehry, Raja Shanker, Syed Zaidi, Fahad AlGhmlas, Ibraheem Motabi, Shahid Iqbal, Ahmad Butt, Hassan AlShehri, Imran Tailor, Syed Altaf, Mubarak AlGhamdi, Mohammed Marie, Mansour AlFayez, Kamal Al Zahrani, Mohammed Dwaimah, Tahani Al-Halouli, Wafaa Al-Shakweer, Maied AlShehery, Abdul Zaidi, Atta Gill, Belal Albutoosh, Musab Ahmed.	627
Implementation Outcome Scales for Digital Mental Health (iOSDMH): Scale Development and Cross-sectional Study (e24332)	
Natsu Sasaki, Erika Obikane, Rajesh Vedanthan, Kotaro Imamura, Pim Cuijpers, Taichi Shimazu, Masamitsu Kamada, Norito Kawakami, Daisuke Nishi.	645
Digital Instruments for Reporting of Gastrointestinal Symptoms in Clinical Trials: Comparison of End-of-Day Diaries Versus the Experience Sampling Method (e31678)	
Abraham Beckers, Johanna Snijkers, Zsa Weerts, Lisa Vork, Tim Klaassen, Fabienne Smeets, Ad Masclee, Daniel Keszthelyi.	657
Limited Interaction Targeted Epidemiology of HIV in Sexual and Gender Minority American Adolescents and Adults: Feasibility of the Keeping it LITE Study (e30761)	
Neil Gleason, Pedro Serrano, Alejandro Muñoz, Audrey French, Sybil Hosek.	685
Twitter Use by Academic Nuclear Medicine Programs: Pilot Content Analysis Study (e24448)	
Ananya Panda, Akash Sharma, Ayca Dundar, Ann Packard, Lee Aase, Amy Kotsenas, Ayse Kendi.	712

Googling for Suicide—Content and Quality Analysis of Suicide-Related Websites: Thematic Analysis (e29146)
 Wen Chen, Andrea Boggero, Giovanni Del Puente, Martina Olcese, Davide Prestia, Haitham Jahrami, Nasr Chalhaf, Noomen Guelmami, Fairouz Azaiez, Nicola Bragazzi. 719

Conducting Health Literacy Research With Hard-to-Reach Regional Culturally and Linguistically Diverse Populations: Evaluation Study of Recruitment and Retention Methods Before and During COVID-19 (e26136)
 Genevieve Perrins, Tabassum Ferdous, Dawn Hay, Bobby Harreveld, Kerry Reid-Searl. 727

Adapting a Mental Health Intervention for Adolescents During the COVID-19 Pandemic: Web-Based Synchronous Focus Group Study (e30293)
 Maria Calvo-Valderrama, Arturo Marroquín-Rivera, Erin Burn, Laura Ospina-Pinillos, Victoria Bird, Carlos Gómez-Restrepo. 743

COVID-19 Knowledge, Attitudes, and Practices Among People in Bangladesh: Telephone-Based Cross-sectional Survey (e28344)
 Md Rabbani, Orin Akter, Md Hasan, Nandeeta Samad, Shehrin Mahmood, Taufique Joarder. 753

Exploring Online Health Reviews to Monitor COVID-19 Public Health Responses in Alabama State Department of Corrections: Case Example (e32591)
 Pamela Valera, David Carmona, Sarah Malarkey, Noah Sinangil, Madelyn Owens, Asia Lefebvre. 774

Early Report

A Program to Improve Digital Access and Literacy Among Community Stakeholders: Cohort Study (e30605)
 Brittany Drazich, Yeukai Nyikadzino, Kelly Gleason. 270

Reviews

Toward Automated Data Extraction According to Tabular Data Structure: Cross-sectional Pilot Survey of the Comparative Clinical Literature (e33124)
 Karl Holub, Nicole Hardy, Kevin Kallmes. 670

Usability-In-Place—Remote Usability Testing Methods for Homebound Older Adults: Rapid Literature Review (e26181)
 Jordan Hill, Janetta Brown, Noll Campbell, Richard Holden. 703

Viewpoint

A Case Study of an SMS Text Message Community Panel Survey and Its Potential for Use During the COVID-19 Pandemic (e28929)
 Lilian Chan, Nouhad El-Haddad, Becky Freeman, Blythe O'Hara, Lisa Woodland, Ben Harris-Roxas. 737

Corrigenda and Addendas

Correction: The Potential of Digital Data Collection Tools for Long-lasting Insecticide-Treated Net Mass Campaigns in Nigeria: Formative Study (e34526)
 Youngji Jo, Nathan Barthel, Elizabeth Stierman, Kathryn Clifton, Esther Pak, Sonachi Ezeiru, Diwe Ekweremadu, Nnaemeka Onugu, Zainab Ali, Elijah Egwu, Ochayi Akoh, Orkan Uzunyayla, Suzanne Van Hulle. 784



Correction: Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland ([e34897](#))

Kristjana Thorarinsdottir, Emily Holmes, Johann Hardarson, Unnur Hedinsdottir, Marie Kanstrup, Laura Singh, Arna Hauksdottir, Thorhildur Halldorsdottir, Berglind Gudmundsdottir, Unnur Valdimarsdottir, Edda Thordardottir, Beau Gamble, Andri Bjornsson. 786

Original Paper

A Transdiagnostic Self-management Web-Based App for Sleep Disturbance in Adolescents and Young Adults: Feasibility and Acceptability Study

Nicole E Carmona¹, MA; Aleksandra Usyatynsky¹, MA; Samlau Kutana², BA; Penny Corkum^{3,4}, PhD; Joanna Henderson^{5,6}, PhD; Kelly McShane^{1,7}, PhD; Colin Shapiro^{5,8}, PhD; Souraya Sidani⁹, PhD; Jennifer Stinson^{10,11,12}, PhD; Colleen E Carney¹, PhD

¹Department of Psychology, Ryerson University, Toronto, ON, Canada

²Department of Psychology, Memorial University of Newfoundland, St. John's, NL, Canada

³Department of Psychology and Neuroscience, Dalhousie University, Halifax, NS, Canada

⁴Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

⁵Department of Psychiatry, University of Toronto, Toronto, ON, Canada

⁶Margaret and Wallace McCain Centre for Child, Youth and Family Mental Health, Centre for Addiction and Mental Health, Toronto, ON, Canada

⁷Human Resource Management and Organizational Behaviour, Ted Rogers School of Management, Ryerson University, Toronto, ON, Canada

⁸Department of Ophthalmology, University of Toronto, Toronto, ON, Canada

⁹Daphne Cockwell School of Nursing, Ryerson University, Toronto, ON, Canada

¹⁰Chronic Pain Program, Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, ON, Canada

¹¹Child Health Evaluative Sciences, Research Institute, The Hospital for Sick Children, Toronto, ON, Canada

¹²Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Colleen E Carney, PhD

Department of Psychology

Ryerson University

350 Victoria Street

Toronto, ON, M5B 2K3

Canada

Phone: 1 416 979 5000 ext 552177

Email: ccarney@ryerson.ca

Abstract

Background: Sleep disturbance and its daytime sequelae, which comprise complex, transdiagnostic sleep problems, are pervasive problems in adolescents and young adults (AYAs) and are associated with negative outcomes. Effective interventions must be both evidence based and individually tailored. Some AYAs prefer self-management and digital approaches. Leveraging these preferences is helpful, given the dearth of AYA treatment providers trained in behavioral sleep medicine. We involved AYAs in the co-design of a behavioral, self-management, transdiagnostic sleep app called DOZE (Delivering Online Zzz's with Empirical Support).

Objective: This study tests the feasibility and acceptability of DOZE in a community AYA sample aged 15-24 years. The secondary objective is to evaluate sleep and related outcomes in this nonclinical sample.

Methods: Participants used DOZE for 4 weeks (2 periods of 2 weeks). They completed sleep diaries, received feedback on their sleep, set goals in identified target areas, and accessed tips to help them achieve their goals. Measures of acceptability and credibility were completed at baseline and end point. Google Analytics was used to understand the patterns of app use to assess feasibility. Participants completed questionnaires assessing fatigue, sleepiness, chronotype, depression, anxiety, and quality of life at baseline and end point.

Results: In total, 83 participants created a DOZE account, and 51 completed the study. During the study, 2659 app sessions took place with an average duration of 3:02 minutes. AYAs tracked most days in period 1 (mean 10.52, SD 4.87) and period 2 (mean 9.81, SD 6.65), with a modal time of 9 AM (within 2 hours of waking). DOZE was appraised as highly acceptable (mode \geq 4) on the items "easy to use," "easy to understand," "time commitment," and "overall satisfaction" and was rated as credible (mode \geq 4)

at baseline and end point across all items (logic, confident it would work, confident recommending it to a friend, willingness to undergo, and perceived success in treating others). The most common goals set were decreasing schedule variability (34/83, 41% of participants), naps (17/83, 20%), and morning lingering in bed (16/83, 19%). AYAs accessed tips on difficulty winding down (24/83, 29% of participants), being a night owl (17/83, 20%), difficulty getting up (13/83, 16%), and fatigue (13/83, 16%). There were significant improvements in morning lingering in bed ($P=.03$); total wake time ($P=.02$); sleep efficiency ($P=.002$); total sleep time ($P=.03$); and self-reported insomnia severity ($P=.001$), anxiety ($P=.002$), depression ($P=.004$), and energy ($P=.01$).

Conclusions: Our results support the feasibility, acceptability, credibility, and preliminary efficacy of DOZE. AYAs are able to set and achieve goals based on tailored feedback on their sleep habits, which is consistent with research suggesting that AYAs prefer autonomy in their health care choices and produce good results when given tools that support their autonomy.

Trial Registration: ClinicalTrials.gov NCT03960294; <https://clinicaltrials.gov/ct2/show/NCT03960294>

(*JMIR Form Res* 2021;5(11):e25392) doi:[10.2196/25392](https://doi.org/10.2196/25392)

KEYWORDS

youth; sleep; technology; mHealth; self-management; adolescents; young adults; mobile phone

Introduction

Background

Emerging adulthood is a time of risk for sleep problems. An epidemic of shortened sleep times and resultant sleepiness in teens and young adults has been reported worldwide [1,2]. Sleep problems are reported in two-thirds of adolescents and young adults (AYAs) [3-5] and are linked to academic and mental health problems and suicidality [6-11]. AYAs also experience a pubertal shift toward later or delayed circadian rhythmicity, such that they become tired later, resulting in later rise times, although they must maintain early rise times to attend school [12-15]. This biological shift toward a later sleep onset time is further compounded by the fact that AYAs are sensitive to evening light [12,16,17] and are exposed to academic (eg, early school start times) and social pressures (eg, peers reinforcing later activities), which can increase sleeplessness. Moreover, some AYAs exert control over what they eat and drink (eg, caffeine) without an adequate understanding of the consequences of particular behaviors on their sleep and general health [18]. As a result, AYAs have chronic sleep loss, resulting in pathological sleepiness, a condition associated with an increased risk of motor vehicle accidents and poor emotional, interpersonal, and academic functioning [19,20].

Providing accessible interventions for sleep problems is crucial to mitigate personal, social, and economic costs. Treatment options for most sleep problems typically consist of pharmacological and behavioral approaches [21,22]. Sleep medications are associated with significant concerns regarding use by AYAs [23]. Although cognitive behavioral therapy for insomnia (CBT-I) approaches are the most efficacious for helping individuals with insomnia with few troublesome side effects [24], few AYAs have access to behavioral sleep medicine treatments [25,26]. Furthermore, CBT-I protocols (when accessible) do not target the range of sleep problems that AYAs experience beyond insomnia symptoms, including hypersomnia, excessive daytime sleepiness (eg, from voluntary sleep restriction), delayed phase circadian rhythm, and often uninformed use of sleep-interfering substances, some of which may be experienced concurrently, highlighting the promise of transdiagnostic behavioral sleep treatments for AYAs [27-29]. There are multiple barriers to accessing nonpharmacological

evidence-based sleep treatments, including poor dissemination of CBT-I and behavioral sleep treatments in real-world settings [30] and a shortage of providers who are trained in such treatments [31]. It is clear that AYAs have diverse sleep problems (ie, not just insomnia) and require tailored evidence-based strategies that comprehensively address these diverse needs, making transdiagnostic cognitive behavioral treatments [32] a more suitable option than nonadapted CBT-I.

Smartphone apps are uniquely suited to improve access for those who are not seeking treatment but are nevertheless at risk as well as treatment-seeking groups with poor access to evidence-based treatments, and reportedly increase users' motivation, goal-setting, and confidence in and control over their ability to enact health-promoting behavior change [33]. AYAs are drawn to technology; therefore, technology provides an opportunity to directly reach this large invisible at-risk cohort for whom sleep disorders contribute to psychopathology and vice versa [34,35]. Although there are web-based programs for infant (aged 6-36 months) [36], pediatric (aged 1-8 years) [37,38], adolescent [39-41], and adult sleep disturbances [42], there are no evidence-based transdiagnostic programs that adequately address the *multitude* of sleep problems observed in AYAs and that span both adolescence and young adulthood. Put differently, as some of the sleep problems of AYAs are unique and go beyond insomnia (eg, voluntary sleep restriction, circadian phase delay, and poor sleep hygiene), pediatric and adult programs are not suitable for this age group.

Objective

In response to a dearth of appropriate treatment options for AYAs with a range of sleep difficulties beyond insomnia, we created a self-management app called DOZE (Delivering Online Zzz's with Empirical Support), which was co-designed by AYAs in user consultation sessions conducted by the developers of the app. DOZE is based on the principles of cognitive behavioral sleep medicine [32] and allows AYAs access to effective, tailored strategies to address a range of sleep problems (eg, insomnia, daytime sleepiness, and delayed phase circadian rhythms), which also range from subclinical to clinical in terms of their severity. Importantly, consistent with their preferences for technology and autonomy and as revealed by the user consultation sessions, the self-management approach entrusts

AYAs to set their own goals and make their own health behavior changes when provided with the tools and education to do so. This study reports on an open trial evaluating the app's feasibility over a 4-week intervention period. The primary objective of this study is to evaluate whether AYAs regard evidence-based cognitive behavioral strategies in a self-management app (DOZE) as acceptable and credible. Toward the aim of evaluating feasibility, we also want to understand how AYAs use DOZE and which aspects of the app they find most valuable. As AYAs co-designed the app, we hypothesize that AYAs would rate the app as acceptable and credible and report high satisfaction. No a priori hypotheses are made regarding which aspects of the app would be most valuable to AYAs, and an exploratory analysis of app use patterns is conducted to understand which aspects of the app are most useful and to provide further insight into AYAs' primary sleep complaints and treatment goals.

The secondary objective is to examine preliminary efficacy, acknowledging that these results may be quite modest because of the range of sleep problems addressed (eg, both insomnia and hypersomnia) and that, without sleep disturbance severity as part of the inclusion criteria, we might recruit a subclinical sample. Outcomes focus on whether DOZE positively affects health-related behavior change, improvements in daytime symptoms (ie, sleepiness, fatigue, energy, and psychological symptoms), and health-related quality of life. We hypothesize that AYAs who use DOZE would see improvements in sleep indices at end point. Furthermore, we hypothesize that the use of DOZE for 4 weeks may contribute to improvements in energy, mood, and health-related quality of life at end point compared with baseline.

Methods

Participants

Individuals between the ages of 15 and 24 years with self-reported dissatisfaction with their sleep were eligible to participate in this web-based study. Given that a clinical diagnosis of a sleep disorder is not required to use DOZE (ie, any AYA who wants to improve sleep can use it), this study did not have any other inclusion or exclusion criteria to be able to generalize the results of this study to the population that will use DOZE. Self-reported dissatisfaction with sleep was queried with a single question, "How do you feel about your sleep?" and participants could respond using an open textbox. Computer and internet literacy of participants was assumed. Participants were recruited using posters at Ryerson University and in the downtown Toronto community, dissemination of study information to mental health providers via listservs of professional organizations (eg, Canadian Association for Cognitive and Behavioural Therapy and Ontario Psychological Association), and via social media (eg, Twitter and Facebook). Recruitment materials advertised the opportunity for AYAs who were dissatisfied with their sleep or were experiencing difficulty sleeping to participate in a study testing a new sleep app for 4 weeks. Participation in this study was quasi-anonymous. Email confirmation was used to detect repeat participation. All participants provided informed consent and

completed a quiz assessing their understanding of critical components of the consent form (detailed in the *Procedure* section).

Intervention: DOZE App

DOZE [43] is a transdiagnostic, self-management, cognitive behavioral web-based app for sleep disturbance among AYAs. It is freely accessed from a web browser on any device (ie, it is not downloaded on a device). DOZE was designed using an iterative, experience-based co-design process that included feedback from AYA users and health care provider stakeholders [44-46]. Initial design consultation interviews conducted by an industry partner with AYA stakeholders revealed that, in addition to including a sleep diary for self-monitoring, AYAs wanted an app that provides feedback on their personal sleep habits (over and above feedback on their sleep), opportunities for goal-setting, and tips on how to make changes to their sleep. Following this initial phase, a user-informed redesign of DOZE was conducted with an industry partner (PIVOT Design), consistent with standard app development processes [47], including 6 user assessments with AYAs (aged 15-24 years) on iterative prototype versions of the app and consultations with health care provider stakeholders. Of note, AYAs who contributed to the co-design of DOZE were not participants in this study. Additional details about the co-design process are presented in [Multimedia Appendix 1](#).

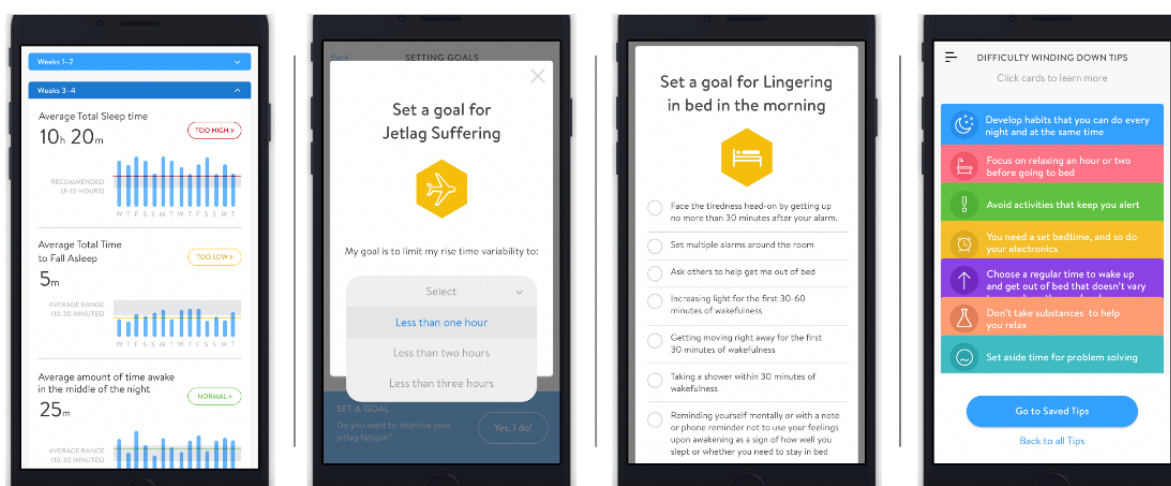
The final version of DOZE includes a sleep diary, adapted with permission from the Consensus Sleep Diary [48], comprising 8 questions that assess the following: bedtime (ie, time of getting into bed whether or not it was to initiate sleep), time of sleep attempt (ie, time of trying to sleep), time it took to fall asleep (sleep onset latency [SOL]), total time awake in the night (wake after sleep onset), time of final awakening (ie, time of waking in the morning), rise time (ie, time of getting out of bed), amount of time spent napping, and use of sleep-interfering substances within 2 hours of bedtime. DOZE's onboarding screens instructed participants to complete their sleep diary every morning within 2 hours of waking, and they were reminded of this using a pop-up notification if they completed it that same day outside of the 2-hour window. Completing an entry for the previous day was not permitted. Participants also received an automated reminder email each morning to complete the diary. Sleep indices were calculated based on information provided by the sleep diary, including SOL, variability in rise time (range of earliest to latest rise time), variability in bedtime (range of earliest to latest bedtime), wake after sleep onset, morning lingering in bed (the difference between time of final awakening and rise time), total sleep time (TST; total time spent asleep), total wake time (TWT; total time spent awake during the sleep period), time in bed (TIB; bedtime to rise time), and sleep efficiency (SE; percent of TIB spent asleep). Participants were able to view their average sleep indices over a 2-week period on a personalized dashboard.

After completing 2 weeks of sleep diaries, DOZE provides personalized feedback on sleep indices (TST, SOL, TWT, TIB, and SE) relative to age-adjusted norms [4] by indicating whether their indices are normal, too high, or too low on the Progress screen. Users then have the opportunity to set goals to improve

their sleep in *problem areas* that are identified (ie, indices are outside the normal range). Feedback areas, which all represent components of evidence-based treatment [32], include *naps*, *too much/too little TIB*, *sleep-interfering substances*, *lingering in bed in the morning*, *sleepiness*, and *jetlag without traveling* (ie, sleep schedule variability). AYAs have the opportunity to set personalized goals in these areas (eg, reduce jetlag by 1 hour, 2 hours, or 3 hours), which are added to their personal dashboard, or they may choose not to set goals in identified areas, accounting for various stages of readiness to change. Notably, if a feedback area is not relevant to the individual based on their sleep indices (eg, they are spending a normal amount of TIB), they are not provided the opportunity to set a goal to improve in that area (eg, they are not shown the feedback area *too much/too little TIB*). In addition, participants could access

a Tips section that included quizzes and psychoeducation about hyperarousal (*difficulty winding down*), chronotype (*night owl living in an early bird's world*), difficulty getting out of bed in the morning, daytime sleepiness (*trouble staying awake during the day*), and fatigue (*exhausted*). Tips can also be saved to participants' dashboards for a personalized user experience, and all participants have access to all of the tips. After setting goals, participants then complete sleep diaries for an additional 2 weeks and, at the conclusion of the second 2-week period, they are able to see a comparison of their progress between weeks 1-2 and weeks 3-4 on their dashboard. Screenshots of the example progress, goal-setting, and tips screens are presented in [Figure 1](#). Participants did not receive any other support or interventions in this study.

Figure 1. Screenshots of Delivering Online Zzz's with Empirical Support: progress and feedback, goal-setting, and tips.



Measures

Outcomes were assessed using analysis of aggregate DOZE use during the study period and through the completion of web-based qualitative and quantitative questionnaires.

Feasibility and Acceptability

Treatment acceptability, credibility, and satisfaction were assessed using both quantitative and qualitative measures. A 5-item Treatment Evaluation Questionnaire (TEQ) was administered at baseline and end point and assessed participants' perceptions of DOZE as a logical treatment, beliefs that it would be successful, confidence in referring it to a friend, willingness to undergo treatment, and whether they thought it was successful in treating others [49]. As reviewed in the *Introduction* section, DOZE is a self-management intervention rather than a treatment for clinical sleep disorders. The use of the word *treatment* in evaluating the perceived credibility of DOZE reflects the verbatim wording of each item of the TEQ, which has been used in previous clinical trials of sleep medicine treatments [50]. Each item is scored from 1 (*not at all*) to 7 (*very*), with higher scores reflecting greater perceived credibility and satisfaction with DOZE. In addition, a 6-item Acceptability Scale was

designed for the purpose of this study and administered at end point. This measure assessed participants' attitudes regarding ease of use, understandability, enjoyment, perceived helpfulness, time commitment, and overall satisfaction. Participants rated each item on a scale of 1-5, with higher scores denoting greater acceptability and more favorable attitudes toward DOZE. In addition, participants completed a qualitative exit questionnaire about their experience using DOZE at end point, including how they felt about their sleep and the goals they set, what they liked and did not like, and whether there was anything they would change about DOZE. An analysis of qualitative responses will be reported in a separate paper.

Google Analytics was used to provide additional information about DOZE use, including the number of page views; where, when, and how DOZE was accessed; and the mean length of sessions using DOZE. Consistent with what is available on the Google Analytics website, only averages are provided, as other measures of central tendency are unavailable.

Preliminary Efficacy

Improvements in sleep and sleep-related behaviors were evaluated by analyzing changes in the following sleep and

behavioral indices derived from the sleep diary: variability in bedtime, variability in sleep attempt, variability in wake time, variability in rise time, morning lingering in bed, SE, TST, TWT, TIB, duration of naps, and frequency of sleep-interfering substance use. In addition, participants completed self-report questionnaires at baseline and end point to investigate whether the use of DOZE for 4 weeks contributed to improvements in insomnia severity, chronotype, daytime symptoms, and health-related quality of life.

The Insomnia Severity Index (ISI [51]) is a 7-item scale that evaluates self-reported sleep disturbances. Items are scored on a 5-point Likert scale from 0 (*no sleep difficulty*) to 4 (*severe sleep difficulty*), with total scores ranging from 0 to 28. A cutoff score of 10 distinguishes between good and poor sleepers in adult community samples [52]. Previous research has suggested a cutoff score of 9 among adolescent samples [53]; however, because of the larger proportion of young adults in our sample relative to adolescents (see the *User Characteristics* section), we used the more conservative cutoff score of 10 with the understanding that clinical levels of insomnia may be underestimated in our sample. The ISI has been validated in AYA samples and is a validated outcome measure with good psychometric properties [53-56].

Chronotype, which refers to one's endogenous circadian preference, was measured using the 13-item Composite Scale of Morningness (CSM [57]). Total scores range from 13 to 55, with higher scores reflecting greater morningness preference; based on established cutoffs, scores between 44 and 55 indicate morning type, between 23 and 43 indicate intermediate type, and between 13 and 22 reflect evening type [57]. The CSM has been validated in samples of adolescents and undergraduate students [57-59].

Fatigue was measured using the Fatigue Severity Scale (FSS [60]). The FSS is a 9-item scale measuring impairment related to fatigue, and items are rated on a Likert scale from 1 (*no impairment*) to 7 (*severe impairment*). A mean item score is calculated, with higher scores indicating more severe fatigue. The FSS has demonstrated convergent validity with a visual analog scale for fatigue [60] and has previously been used to measure fatigue in clinical samples of adolescents [61,62].

Daytime sleepiness was measured using the Cleveland Adolescent Sleepiness Questionnaire (CASQ) [63], which comprises 16 items scored on a 5-point Likert scale ranging from *never* to *almost every day*. Items are scored to yield a total score, as well as a Sleepiness subscale and a reverse-coded Alertness subscale, with higher scores indicating greater daytime sleepiness. The CASQ has been validated in samples of clinical and nonclinical adolescents [63].

The Center for Epidemiologic Studies Depression Scale–Revised 10-item Version for Adolescents (CESDR-10 [64]) was included to evaluate depression symptom severity. The item assessing suicidality was removed to mitigate risk (because of the nature of the web-based study), leaving a 9-item scale. Participants rated the frequency of each symptom during the past 2 weeks, ranging from 0 (*not at all or less than 1 day*) to 4 (*nearly every day for 2 weeks*). In this study, total scores ranged from 0 to 36, with higher scores reflecting greater depression symptom

severity. The CESDR-10 has demonstrated excellent internal consistency in adolescents and is a valid screening measure for depression [64].

Anxiety symptoms were assessed using the State-Trait Inventory of Cognitive and Somatic Anxiety (STICSA [65]), which comprises scales assessing the presence of cognitive and somatic symptoms of anxiety at the moment it is being completed (State) and in general (Trait). The State and Trait scales include 21 statements, and agreement is rated on a scale from 1 (*not at all*) to 4 (*very much so*). Total scores on both scales range from 21 to 84. The STICSA has demonstrated reliability and validity in undergraduate and adolescent samples [66-68] and has demonstrated sensitivity to change [68].

The RAND 36-item Short Form Health Survey 1.0 (SF-36 [69]) was included in this study to evaluate health-related quality of life. The SF-36 consists of 36 items that assess the following domains (scales): physical functioning, role functioning–physical, role functioning–emotional, energy, emotional well-being, social functioning, pain, and general health. Each scale score ranges from 0-110, with higher scores denoting better health functioning. Each scale has demonstrated adequate to excellent internal consistency [70].

Procedure

Interested AYAs were directed to the DOZE website, where they were provided with a consent form and details of our privacy policy. As the study took place on the web, a brief quiz followed the consent form to ensure that participants understood key aspects of informed consent, including the purpose of the study, risks and benefits of participation, and limits of confidentiality. After the provision of informed consent, participants received a randomly generated 5-digit study ID and completed a battery of questionnaires using Qualtrics, which included a demographic form and the ISI, FSS, CSM, CASQ, CESDR-10, STICSA State and Trait, SF-36, and TEQ. At the end of the questionnaires, participants received a link to access DOZE and create an account.

As indicated by the in-app onboarding process, participants were instructed to use DOZE for 2 weeks (ie, period 1), after which they would receive personalized feedback on their sleep, have the opportunity to set goals and access tips, and track their sleep for 2 more weeks (ie, period 2) to examine their progress. Participants were contacted via email by the study coordinator to maintain adherence with sleep tracking if they missed several consecutive days of completing the sleep diary and were sent a standardized email encouraging them to get back on track the next day. Following the completion of the 4-week study period, participants were emailed a link to complete the web-based posttest battery of questionnaires that included the ISI, FSS, CSM, CASQ, CESDR-10, STICSA State and Trait, SF-36, and TEQ, as well as the Acceptability Scale and the Qualitative Exit Questionnaire. Participants were compensated CAD \$12.50 (US \$9.79) for each study component that they completed; thus, participants who completed all 4 components of the study were emailed a CAD \$50 (US \$39.17) gift card honorarium for their participation. This study was approved by the research ethics board of Ryerson University, Toronto, Canada. Data collection took place between September 2019 and January 2020. There

were no significant changes to the methods of the study or the intervention (eg, bug fixes, downtimes, or content changes) while this study was underway.

Statistical Analysis

User characteristics, including self-reported demographic and baseline clinical characteristics as well as sleep indices from period 1, were evaluated using descriptive statistics for all participants who created a DOZE account. The demographic characteristics and scores on all self-report measures at baseline were compared between early dropouts (ie, before, during, and after the baseline questionnaires) and those who went on to create a DOZE account to assess for possible differences, as well as between all dropouts and those who completed the study (*study finishers*). To evaluate the feasibility of DOZE, including the characteristics and app use patterns of all DOZE users, analyses of app use (ie, number of days of sleep diary completion, use of goals and tips, and Google Analytics) included participants who created an app account regardless of whether or not they completed the study. Histograms and P-P plots were generated for visual inspection of normality, and extreme outliers on sleep indices were removed on the basis of visual inspection of boxplots. Descriptive statistics were used to analyze demographic and clinical characteristics. Patterns of app use were analyzed using descriptive statistics and Google Analytics.

Owing to insufficient sleep diary data from DOZE users who did not complete the study (32/83, 39%), sleep diary indices

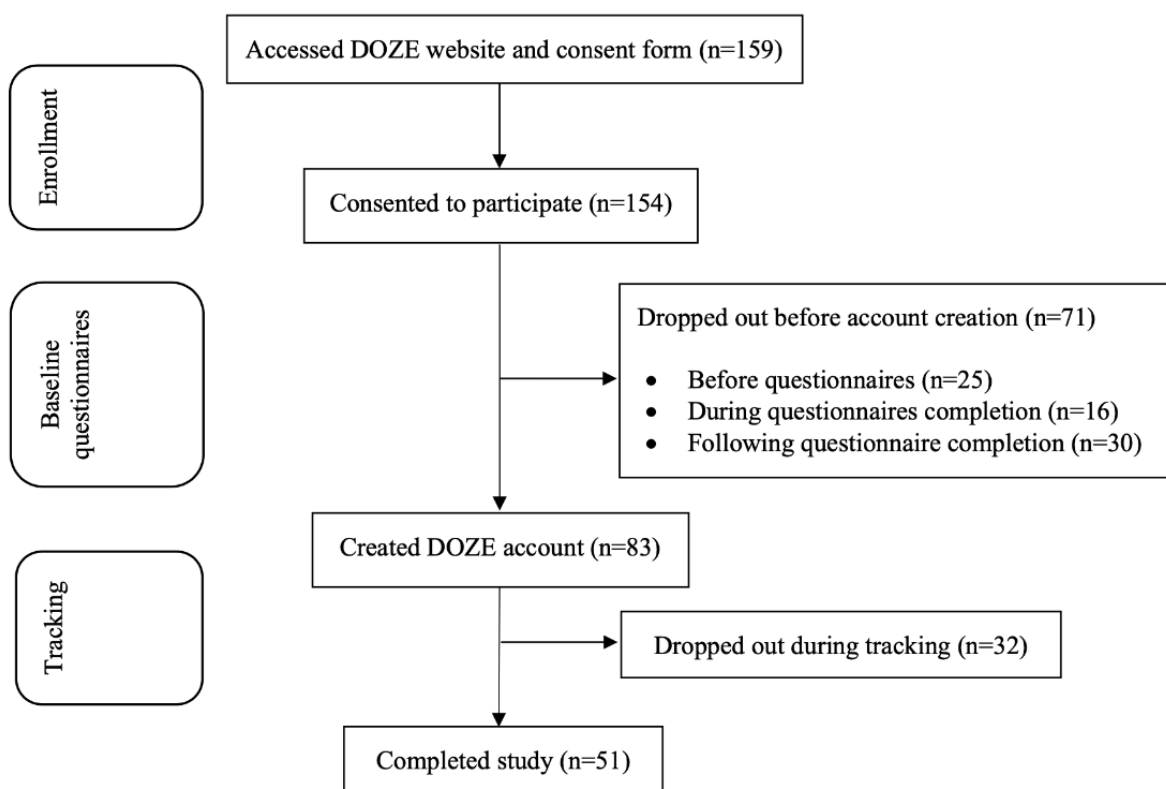
are only presented for study finishers. In addition, analyses evaluating acceptability, credibility, and satisfaction with DOZE, and examining the efficacy of DOZE in improving sleep indices and self-reported measures of sleep and daytime functioning were only conducted among study finishers. Paired sample 2-tailed *t* tests were used to analyze changes in satisfaction with DOZE from baseline to end point, and acceptability was analyzed using descriptive statistics. Changes in sleep indices from period 1 to period 2 were analyzed using paired sample *t* tests. Similarly, paired sample *t* tests with 1000 bootstrapped samples were used to evaluate differences in self-reported insomnia severity, daytime symptoms, and health-related quality of life from baseline to end point to account for slight deviations from normality. Effect sizes were calculated for all analyses of change from baseline to end point using Cohen *d* for repeated measures. Unless otherwise specified, all analyses were performed using SPSS for Mac version 26 (IBM Corporation).

Results

User Characteristics

In total, 154 participants consented to participate in this study (Figure 2). Of the 154 participants who consented to participate, 71 (46.1%) dropped out during the baseline questionnaire process. Over half of the participants (83/154, 53.9%) created an app account and completed some tracking. Approximately 33.1% (51/154) of participants completed all 4 components of the study (baseline questionnaires, first 2 weeks of sleep diaries, second 2 weeks of sleep diaries, and end point questionnaires).

Figure 2. Flowchart of participants. DOZE: Delivering Online Zzz’s with Empirical Support.



Demographic characteristics were analyzed using data from all DOZE users (83/83, 100%; Table 1). Participants were on

average 20.17 years (SD 2.49 years). Out of the 83 participants, 20 (24%) were adolescents (between the ages of 15 and 18

years) and 63 (76%) were young adults (between the ages of 19 and 24 years). The sample was predominantly female (57/83, 69%) and European Canadian (37/83, 45%). Most participants were not currently seeing a health care professional for the purpose of a sleep problem (79/83, 95%) or a mental disorder (69/83, 83%). Despite low levels of treatment-seeking, the average score on the ISI indicated sleep disturbance in the mild range (mean 12.48, SD 4.67), and 76% (63/83) of participants exceeded the clinical cutoff on the ISI (ie, score > 10). In addition, the sample was characterized by high levels of fatigue, as indicated by the FSS (mean 4.73, SD 1.01) and the energy scale of the SF-36 (mean 35.71, SD 18.24) and daytime sleepiness (mean 45.14, SD 9.47) and on average had an intermediate chronotype on the CSM (mean 30.07, SD 7.16). Levels of state

(mean 42.25, SD 12.73) and trait anxiety (mean 42.71, SD 13.15), as well as depression (mean 16.85, SD 8.30), were mild in this sample. Notably, early dropouts of the study (ie, before, during, and after the baseline questionnaires) did not differ from DOZE users on age ($t_{114}=0.40$; $P=.69$), gender ($\chi^2_2=3.0$; $P=.22$), ethnicity ($\chi^2_9=16.3$; $P=.06$), ISI ($t_{124}=1.17$; $P=.25$), FSS ($t_{124}=1.08$; $P=.28$), CMQ ($t_{118}=0.63$; $P=.53$), CASQ ($t_{113}=-0.60$; $P=.55$), STICSA State ($t_{111}=0.74$; $P=.46$), STICSA Trait ($t_{104}=0.41$; $P=.69$), CESDR-10 ($t_{115}=-0.03$; $P=.98$), or any of the SF-36 scales (P values range from .55 to .99). Similarly, study finishers did not differ from participants who dropped out at any point on any of these demographic or clinical characteristics (data not shown).

Table 1. Demographic characteristics of participants at baseline (N=83).

Baseline characteristics	Participants, n (%)
Sex	
Female	57 (69)
Male	26 (31)
Ethnicity	
Caribbean Canadian	1 (1)
East or Southeast Asian Canadian	15 (18)
European Canadian	37 (45)
Latin American, Central American, or South American Canadian	4 (5)
South Asian Canadian	8 (10)
West Asian or Arab Canadian	1 (1)
Pacific Islander Canadian	1 (1)
Other	16 (19)
Seeing an HCP ^a for sleep disorder ^b	4 (5)
Seeing an HCP for mental disorder ^b	14 (17)

^aHCP: health care professional.

^bReflects the number and percentage of participants who answered "yes" when asked about whether they were currently seeing a health care professional for this issue.

We analyzed sleep indices for study finishers at periods 1 and 2 (Table 2). After removing outliers (ranging from n=1 to n=6 depending on the sleep index), period 1 sleep indices indicated that despite an average SE in the normal range (ie, >85%), DOZE users spent a clinically meaningful amount of time awake at night. DOZE users also had variable sleep schedules, characterized by >4 hours of variability in their bedtime, sleep

attempt, wake time, and rise time over a 2-week period. In addition, DOZE users obtained 7 hours 20 minutes of sleep on average despite spending 8 hours 21 minutes in bed; a one-sample t test indicated that this is significantly lower than 8.5, the minimum recommended amount of TST for AYAs ($t_{50}=-8.76$; $P<.001$). In addition, DOZE users spent an average of nearly half an hour lingering in bed in the morning.

Table 2. Change in sleep indices for DOZE (Delivering Online Zzz's with Empirical Support) study finishers^a.

Sleep diary variable	Period 1, mean (SD)	Period 2, mean (SD)	<i>t</i> test ^b (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>
Bedtime variability (hours)	4.26 (2.16)	4.96 (4.08)	0.52 (49)	.60	0.07
Sleep attempt variability (hours)	4.04 (2.19)	4.01 (2.13)	0.88 (49)	.38	0.12
Wake time variability (hours)	4.59 (2.47)	4.62 (2.38)	-0.22 (49)	.83	0.03
Rise time variability (hours)	4.71 (2.41)	5.05 (2.83)	-0.13 (49)	.90	0.02
Nap duration (hours)	0.370 (0.50)	0.220 (0.32)	1.63 (32)	.11	0.28
Morning lingering in bed (hours)	0.402 (0.28)	0.353 (0.30)	2.24 (46)	.03	0.30
Time in bed (hours)	8.35 (1.06)	8.49 (0.85)	-2.40 (49)	.02	0.34
Total sleep time (hours)	7.34 (0.93)	7.67 (0.97)	-4.23 (49)	<.001	0.60
Total wake time (hours)	1.00 (0.54)	0.837 (0.55)	2.54 (47)	.01	0.36
Sleep efficiency (%)	87.67 (6.44)	90.18 (6.46)	-3.43 (49)	.001	0.48
Frequency of sleep-interfering substance use ^c	0.142 (0.21)	0.131 (0.21)	0.34 (49) ^d	.74	0.05

^aTest statistic and Cohen *d* are for log-transformed variables, whereas means and SDs refer to untransformed variables.

^bTest is 2-tailed.

^cFrequency of substance use refers to the proportion of days within a period in which sleep-interfering substance use was endorsed in the sleep diary.

^dNot log-transformed.

Feasibility and Acceptability

Feasibility

Among all participants who created an app account, sleep diaries were completed for an average of 10.52 days (SD 4.87 days; range 0-14 days; 82/83, 99%) in period 1 and 9.81 days (SD 6.65 days; range 0-2 days; 75/83, 90%) in period 2. Over two-third of app users (57/83, 69%) used DOZE long enough to receive feedback on their sleep and to set goals; of these 57

DOZE users, 46 (81%) chose to set one or more goals to improve their sleep. Participants set an average of 1.54 goals (SD 1.12; range 0-4 goals; 57/83, 69%) following period 1.

The most frequent goal was to reduce schedule variability (jetlag), followed by reducing naps and reducing lingering in bed in the morning (Table 3). The most frequent tips that participants accessed to help with their goals were how to wind down before bed and how to help with being a night owl (Table 4).

Table 3. Frequency of goal use by participants (N=83).

Goal	Participants, n (%)
Reduce schedule variability (jet lag)	34 (41)
Reduce to <1 hour	11 (13)
Reduce to <2 hours	18 (22)
Reduce to <3 hours	5 (6)
Limit time in bed	11 (13)
Spend 8.5-10.5 hours on weekdays	4 (5)
Spend 8.5-10.5 hours on both weekend days and weekdays	7 (8)
Manage sleepiness	4 (5)
Target lingering in bed^a	16 (19)
Limit lingering to 30 minutes after alarm	11 (13)
Set multiple alarms around room	1 (1)
Ask others to help them get up	0 (0)
Increase light exposure for first 30-60 minutes	3 (4)
Get moving for the first 30 minutes	5 (6)
Take a shower within first 30 minutes	4 (5)
Remind self not to rely on initial feelings upon waking	5 (6)
Reduce naps	17 (20)
No naps	12 (14)
Decrease number of naps	1 (1)
Decrease length of naps	4 (5)

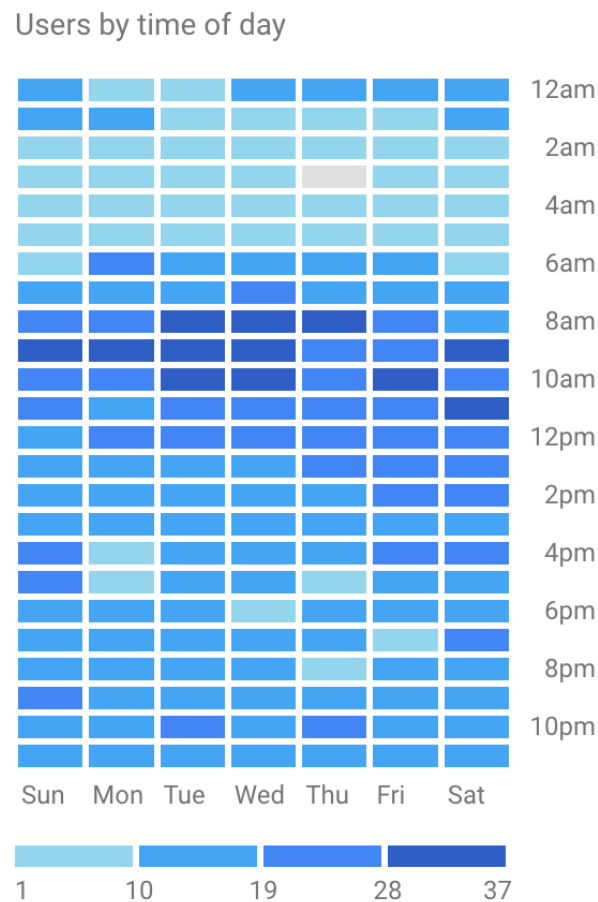
^aParticipants were free to select as many strategies as they desired from a drop-down list.

Table 4. Frequency of tip use by participants (N=83).

Tip access	Participant, n (%)
Completed night owl quiz	17 (20)
Completed difficulty winding down quiz	24 (29)
Completed difficulty getting up quiz	13 (16)
Completed sleep drunkenness quiz	0 (0)
Completed trouble staying awake quiz	5 (6)
Completed fatigue quiz	13 (16)

Google Analytics indicated that 2659 sessions (ie, instances of app use) took place. The average session duration was 3:02 minutes. During the study period, there were 27,263 page views. On average, 10.91 pages were viewed per session (including multiple visits to the same page). When grouped according to content categories, pages related to the sleep diary accounted for most of the 27,263 page views (18,018/27,263, 66.09%), indicating that it was the most used component of the app, followed by Progress (4098/27,263, 15.03%), Dashboard (1925/27,263, 7.06%), Tips (1061/27,263, 3.89%), Goals (810/27,263, 2.97%), Log-In (777/27,263, 2.85%), onboarding

screens (431/27,263, 1.58%), and visits to their user profile (142/27,263, 0.52%). The average time on a page was 18.15 seconds across all pages, ranging from 65.51 seconds for progress pages to 6.68 seconds for onboarding pages. DOZE was most frequently accessed using a mobile device (2,705/2659 sessions, 78.04%), followed by a desktop computer (546/2,659 sessions, 20.53%) or a tablet (38/2,659 sessions, 1.43%). Consistent with the sleep diary and in-app recommendations to complete the sleep diary within 2 hours of waking up, DOZE was most frequently visited between 8 AM and 11 AM (Figure 3), with a modal time of 9 AM.

Figure 3. Users of Delivering Online Zzz's with Empirical Support by time of day.

Acceptability, Credibility, and Satisfaction

To assess the acceptability, credibility, and satisfaction of DOZE, we analyzed the mean, mode, and range of responses for each item of the TEQ and the Acceptability Scale (Table 5). Responses to the TEQ were favorable at baseline and end point, with mean and modal values exceeding 4 for most items. There was no statistical change from baseline to end point on items assessing logic ($P=.54$), confidence that it would work ($P=.48$), confidence in recommending the app to a friend ($P=.74$), and how successful they predicted it would be in helping others with their sleep ($P=.78$), suggesting that most perceptions of DOZE

as a credible and satisfactory intervention were stable over time. However, there was a statistically significant change in the item assessing willingness ($P=.008$). Responses on the Acceptability Scale indicated agreement or high agreement (mode=4 or 5) that DOZE was easy to use, easy to understand, required a reasonable amount of time, and that users were generally satisfied with the app, whereas responses were mostly neutral (mode=3) for enjoyment using the app and how helpful it was in describing their symptoms and quality of life. Averages for each item on the Acceptability Scale ranged from 3.61 (helpful in describing symptoms) to 4.35 (understandable).

Table 5. Participants' responses on the Treatment Evaluation Questionnaire and Acceptability Scale at baseline and during end point assessments (N=51).

Index	Baseline		End point		P value
	Value, mean (SD; range)	Value, mode	Value, mean (SD; range)	Value, mode	
How logical ^a	4.47 (1.72; 1-7)	5	4.71 (1.63; 1-7)	4	.54
Confident in success ^b	4.04 (1.71; 1-7)	4	3.96 (1.48; 1-7)	4	.48
Recommend to friend ^c	4.37 (1.58; 1-7)	5	4.53 (1.55; 1-7)	5	.74
How willing ^d	5.51 (1.73; 1-7)	7	4.49 (1.59; 1-7)	5	.008
How successful for others ^e	4.46 (1.49; 1-7)	4	4.53 (1.41; 1-7)	5	.78
Easy to use	N/A ^f	N/A	4.16 (0.90; 2-5)	5	— ^g
Understandable	N/A	N/A	4.35 (0.89; 2-5)	5	—
Enjoyable	N/A	N/A	3.80 (0.96; 2-5)	3	—
Helpful in describing symptoms and quality of life	N/A	N/A	3.61 (1.08; 2-5)	3	—
Amount of time acceptable	N/A	N/A	4.33 (0.91; 2-5)	5	—
Overall satisfaction	N/A	N/A	3.92 (0.82; 2-5)	4	—

^aHow logical does the treatment you are receiving seem to you?

^bHow confident are you that this treatment will be successful in eliminating your insomnia?

^cHow confident would you be in recommending this treatment to a friend?

^dHow willing are you (were you) to undergo this treatment?

^eHow successful do you think this treatment is for treating other people with insomnia?

^fN/A: not applicable; scale not administered at baseline.

^gSignificance testing not available.

Preliminary Efficacy

Sleep Indices

To assess preliminary efficacy, we compared sleep indices at periods 1 and 2 among study finishers (Table 2). The sleep indices were positively skewed and were therefore log-transformed. They were then re-examined to confirm that the assumption of normality was met (all variables were approximately normal) and subsequently analyzed using paired samples *t* tests, with the exclusion of frequency of sleep-interfering substance use, which was not transformed because of the presence of 0 values. Significant improvements of small-to-medium effect sizes were seen for log-transformed morning lingering ($P=.03$), TIB ($P=.02$), TST ($P<.001$), SE ($P=.001$), and TWT ($P=.01$), whereas variability in sleep schedule (ie, bedtime, sleep attempt, wake time, rise time), naps,

and frequency of sleep-interfering substance use did not significantly change (Table 2).

Self-report Measures

Finally, we compared self-report measures at baseline and end point among study finishers (Table 6). To account for slight deviations from normality among some of the variables, paired *t* tests were conducted using bootstrapping. Mean differences (MDs) were resampled 1000 times, and 95% CIs were generated. Statistically significant improvements were seen on the ISI (MD 3.92, SE 0.61; 95% CI 2.82-5.17), CSM (MD -2.15, SE 0.64; 95% CI -3.41 to -0.96), STICSA State (MD 2.36, SE 1.13; 95% CI 0.16-4.64), STICSA Trait (MD 3.97, SE 0.97; 95% CI 2.03-5.87), CESDR-10 (MD 3.10, SE 0.99; 95% CI 1.20-5.18), and SF-36 energy (MD -6.54, SE 2.48; 95% CI -11.28 to -1.79). Changes from baseline to end point were not significant for the remaining measures of fatigue, sleepiness, or health-related quality of life.

Table 6. Comparison of self-reported measures from baseline to end point for study finishers.

Self-report measure	Baseline		End point		P value	Cohen <i>d</i>
	Score, mean (SD)	Cronbach α	Score, mean (SD)	Cronbach α		
ISI ^a	11.65 (4.35)	.77	8.24 (3.95)	.74	.001	0.90
FSS ^b	4.55 (1.04)	.85	4.35 (1.04)	.86	.07	0.20
CSM ^c	30.80 (7.37)	.86	32.47 (7.79)	.88	.006	0.43
CASQ^d total	44.14 (9.49)	.83	43.14 (11.09)	.89	.50	0.15
Sleepiness	27.42 (7.34)	.81	26.54 (8.49)	.88	.63	0.14
Alertness	16.74 (3.80)	.77	16.34 (4.23)	.81	.86	0.12
CESDR-10 ^e	15.62 (8.34)	.89	12.98 (7.89)	.90	.004	0.44
STICSA^f						
State	39.69 (11.39)	.90	37.83 (13.32)	.94	.05	0.19
Trait	40.80 (11.60)	.91	38.83 (14.25)	.94	.002	0.27
SF-36^g						
Physical functioning	84.22 (22.96)	.93	88.43 (20.70)	.95	.09	0.28
Role functioning-physical	65.20 (39.71)	.85	70.59 (39.28)	.89	.27	0.14
Role functioning-emotional	43.33 (44.29)	.88	42.16 (43.12)	.85	.85	0.03
Energy	38.30 (18.16)	.77	43.92 (19.17)	.75	.01	0.38
Emotional well-being	54.35 (21.27)	.83	56.63 (23.21)	.90	.39	0.13
Social functioning	64.22 (27.62)	.81	67.40 (27.17)	.81	.29	0.14
Pain	77.45 (15.33)	.64	76.23 (18.48)	.77	.73	0.08
General health	62.27 (23.58)	.84	63.95 (22.89)	.77	.08	0.10

^aISI: Insomnia Severity Index.

^bFSS: Fatigue Severity Scale.

^cCSM: Composite Scale of Morningness.

^dCASQ: Cleveland Adolescent Sleepiness Questionnaire.

^eCESDR-10: Center for Epidemiologic Studies Depression Scale–Revised 10-item Version for Adolescents.

^fSTICSA: State-Trait Inventory of Cognitive and Somatic Anxiety.

^gSF-36: RAND 36-item Short Form Health Survey 1.0.

Discussion

Principal Findings

We tested the feasibility, acceptability, and credibility of a self-management behavioral sleep web-based app designed by and for AYAs (aged 15–24 years). The results suggest that AYAs rated the intervention as logical, had confidence that the self-management approach would work, and had confidence in recommending DOZE to a friend. Willingness significantly decreased from baseline to end point, perhaps reflecting different stages of readiness to implement the strategies provided, environmental limitations that can influence their ability to implement strategies (eg, early school rise times and high academic demands), and intrinsic values that are a barrier to engaging in the intervention (eg, valuing sleeping in on weekends over maintaining a regular schedule). However, it is important to note that the TEQ was completed before participants received any detailed information about the intervention (they were only aware that it was a new sleep app

for teens and young adults), representing only naïve expectations of DOZE, whereas at end point they were familiar with the intervention components. DOZE was rated as highly acceptable across indices of ease of use, understandability, time commitment, and satisfaction. Perhaps not surprisingly, ratings about DOZE being enjoyable or succinct in describing their particular symptoms or quality of life issues were acceptable but slightly more modest than ease of use, low time commitment, and overall satisfaction. In support of feasibility, participants engaged with DOZE often (indicated by the number of sessions), most frequently visiting pages pertaining to the sleep diary or their progress and did so using various electronic devices (ie, mobile phones, computers, and tablets). Moreover, although this was not a clinical sample, prospective sleep diary insomnia indices significantly improved from period 1 to period 2 (eg, time spent awake and SE), and the amount of sleep they obtained increased too. AYA users reported significant improvements from baseline to end point in insomnia symptom severity, chronotype, depression, anxiety, and energy. Despite

not recruiting a clinical sample, participants in this study overwhelmingly were not treatment-seekers; however, they evidenced very poor health-related quality of life indices and were above clinical cutoffs for insomnia and fatigue, consistent with what we see in the literature [71].

These results are supportive of self-management. AYAs logged on frequently and completed the diary at the instructed time (ie, within 2 hours of waking up in the morning). Moreover, they worked hard at (1) tracking their sleep (ie, tracking for approximately 10 days in each 14-day period) and (2) setting goals that treatment providers would set (eg, limiting schedule variability and naps and decreasing lingering in bed). They accessed tips to help with their goals, particularly how to wind down before bed, how to help with being a night owl, how to get up regularly in the morning, how to help with fatigue, and how to stay out of bed early in the evening. Thus, we can trust AYAs to set their own goals and make behavioral changes to achieve these goals when they are provided with the tools and education to do so autonomously. Trusting patients to make good choices may not come naturally for providers, but the literature suggests that many patients want autonomy in their health care choices [72] and produce good results when given tools that support their autonomy. As a research group, this was an instructive part of the process, and it was driven by both the literature and collaboration with our AYA advisors and stakeholders throughout the co-design process. Perhaps these findings can inform the tailoring of in-person strategies for AYAs to achieve better outcomes.

Statistically significant differences in sleep indices from period 1 to period 2 seemed unlikely, as we recruited a sample that was not selected for clinical levels of sleep disturbance, and we assumed there would be too little variability in these indices to detect significant effects. Furthermore, detecting a change in variables, such as SOL, naps, or TIB, in a nonclinical or subclinical sample in which some have opposing problems (eg, hypersomnia vs insomnia) is also surprising. For example, those with hypersomnia or sleepiness have SOL values that are too low, and those with insomnia have elevated SOL values; use of DOZE could result in changes in opposite directions and, when averaged across participants, appear to have affected no change at all. These problems can also occur in the same AYA during the same week, for example, sleepiness on some nights and insomnia on others. Similarly, some AYAs are told to eliminate naps, and others are told to increase them, resulting in low variability. Therefore, it is encouraging that insomnia indices, such as lingering in bed in the morning, TWT, TST, TIB, and SE, improved. Although AYAs set goals to decrease the variability in their sleep schedules, the improvement did not reach statistical significance. Perhaps this is why they accessed tips about how to get up in the morning (ie, as it is difficult for them to rise at a regular time), and our plan is to expand tips and resources for AYA users on the DOZE website.

An important implication of this research is that this web-based app may be an important tool to increase access to evidence-based interventions. Some AYAs prefer to use technology for health-related behavior change [34,73,74], and smartphone apps appear to have a high uptake potential; 78.04% used a mobile device for DOZE (2,705/2659 sessions), whereas

only 20% used a desktop computer (546/2,659 sessions, 20.5%) and 1.4% used a tablet (38/2,659 sessions). Incorporating AYAs in the co-design process is also a notable strength of this web-based app and may promote continued uptake. That said, future research will need to determine whether there would be health disparities based on access to technology, Wi-Fi, and geographical location. Clinicians report favorable attitudes toward CBT-I apps [75], suggesting the possibility of widespread access to evidence-based sleep treatments. Given the limited number of expert providers in this unique group (ie, not pediatric but not fully adult either) and the need for sleep solutions at the level of primary care and community mental health clinics, this web-based app holds the promise of increasing access to tailored, age-specific care. We intend to conduct a health economic study on the impact of DOZE in the future.

Limitations

There was a 39% (32/83 participants) dropout rate in the self-management intervention portion of the study, and data were not collected on reasons for discontinuation, although this would certainly be valuable. Although we would prefer a lower rate, the rate was lower than that in other sleep app studies in the literature. For example, digital interventions for insomnia have reported rates >50% [76]. Across adolescent mobile health apps more generally, in a study using an internet-delivered, self-help CBT intervention for chronic pain among adolescents (aged 12-17 years), 52% of participants dropped out of the treatment portion of the study [77]. The dropout rate in our study is perhaps not surprising as they were not treatment-seekers, so if their sleep was not a problem, they might have little motivation except for the CAD \$50 (US \$39.17) compensation for participation. The brief intervention period in this study may also have contributed to lower rates of dropout compared with other studies. In addition, participants in this study were emailed by the research coordinator if they did not complete sleep diaries for several consecutive days; as such, adherence may be lower when disseminated in the real world. A future trial could evaluate whether dropout rates are lower in a clinical sample. Moreover, much of the dropout in this study occurred during the baseline questionnaire period, suggesting that the self-report battery was too burdensome. These questionnaires are not part of the app, so this barrier to participation is not present with regular, real-world app use, and future studies can verify if the app dropout rate is actually lower in real-world app use.

Although appropriate for feasibility trials, the use of an open trial suggests caution when interpreting positive outcomes. Without a control group, we could not ascertain whether the results were specific to the app or the increased sleep attention. In addition, we recruited participants using various strategies (eg, through a university, the community, and mental health providers, which are the intended disseminators of DOZE). Although this could be seen as a strength, given that our sample likely reflects those who will use DOZE in the real world, it may also be a limitation as these could be considered 3 distinct samples. Without knowing more about these samples, we could not test for possible differences. In the same vein, although our intended users are all AYAs, we are cautious about stating that

this would be effective in patients, as this was a community sample; although most participants experienced symptoms that were above clinical cutoffs, they were neither clinically assessed nor were they treatment-seekers. We intend to study the properties of the app in those who are treatment-seeking. Our treatment provider stakeholders suggested that they would be most interested in using the diary as an adjunct to their treatment, much like sleep diary apps in adults [78]. DOZE could be used as a tracking and posttherapy relapse prevention tool, with providers prescribing behavioral change rather than the AYA deciding their own treatment course. Given the strong preferences for self-management in this group, it would be interesting to test outcomes with providers versus self-management.

Our findings may also have been influenced by extraneous factors. First, several items in our measure assessing credibility reference specifically to treatment for insomnia and may not have applied equally to all participants, thereby influencing their responses. Moreover, a small percentage of participants were seeing a health care provider for the purpose of a mental disorder or a sleep disorder, which may have also influenced our results given the bidirectional relationship between sleep

and psychopathology. Finally, we did not query medication use and could not rule out the possibility that some participants were using sleep medication, which may have influenced our results; although low levels of treatment-seeking for a sleep disorder suggests that this is unlikely, and participants enrolled in the trial as they were dissatisfied with their sleep despite current treatments, it nevertheless remains a possibility among a small minority of our participants.

Conclusions

In conclusion, given that there were statistically and clinically significant improvements in a nonclinical sample, it is promising for the tool to be used as designed: as a transdiagnostic tool for young adults to engage in sleep self-management digitally. Even those without sleep disorders were interested and effective in making health-related changes, which can translate into improved sleep habits and potential prevention, as well as potentially improving mental wellness in AYAs. Web-based interventions are an acceptable and effective method of facilitating access to evidence-based care among *invisible groups*, with the ultimate effect of reducing the public health burden.

Acknowledgments

The authors would like to acknowledge the contributions of our adolescent and young adult stakeholders who provided input into the design of DOZE (Delivering Online Zzz's with Empirical Support) and its evaluation, including Sarah Lagler-Clarke, who is currently helping design an uptake strategy. The authors also thank their stakeholders in community mental health, public health, and pediatric sleep medicine, namely Gloria Chaim, MSW; Joanne Figliano-Scott, RN; Janine Robb, PhD; and Shelly Weiss, MD. Finally, the authors acknowledge the in-kind support from the Office of the Vice President of Research and Innovation and the Dean of Arts at Ryerson University; their industry partners at PIVOT Design, including Ian Chalmers, Brenda Little, Dave Brennan, and Melvyn Loa; as well as the Canadian Institutes of Health Research for developing this mechanism for youth mental health initiatives and for funding this project under the eHealth Innovation Partnership Program (#143551).

Conflicts of Interest

CEC discloses the current grant funding from the Canadian government (Canadian Institutes of Health Research). CEC is the creator of DOZE (Delivering Online Zzz's with Empirical Support) and the principal investigator of this study; however, CEC does not receive any royalties for DOZE. The remaining coauthors have nothing to declare.

Multimedia Appendix 1

Delivering Online Zzz's with Empirical Support co-design process.

[[DOC File , 28 KB - formative_v5i11e25392_app1.doc](#)]

References

1. Gradisar M, Gardner G, Dohnt H. Recent worldwide sleep patterns and problems during adolescence: a review and meta-analysis of age, region, and sleep. *Sleep Med* 2011 Feb;12(2):110-118. [doi: [10.1016/j.sleep.2010.11.008](https://doi.org/10.1016/j.sleep.2010.11.008)] [Medline: [21257344](https://pubmed.ncbi.nlm.nih.gov/21257344/)]
2. Owens J, Adolescent Sleep Working Group, Committee on Adolescence. Insufficient sleep in adolescents and young adults: an update on causes and consequences. *Pediatrics* 2014 Sep;134(3):921-932 [FREE Full text] [doi: [10.1542/peds.2014-1696](https://doi.org/10.1542/peds.2014-1696)] [Medline: [25157012](https://pubmed.ncbi.nlm.nih.gov/25157012/)]
3. Hicks RA, Fernandez C, Pellegrini RJ. Striking changes in the sleep satisfaction of university students over the last two decades. *Percept Mot Skills* 2001 Dec;93(3):660. [doi: [10.2466/pms.2001.93.3.660](https://doi.org/10.2466/pms.2001.93.3.660)] [Medline: [11806582](https://pubmed.ncbi.nlm.nih.gov/11806582/)]
4. Hysing M, Pallesen S, Stormark KM, Lundervold AJ, Sivertsen B. Sleep patterns and insomnia among adolescents: a population-based study. *J Sleep Res* 2013 Oct;22(5):549-556 [FREE Full text] [doi: [10.1111/jsr.12055](https://doi.org/10.1111/jsr.12055)] [Medline: [23611716](https://pubmed.ncbi.nlm.nih.gov/23611716/)]
5. Sivertsen B, Vedaa Ø, Harvey AG, Glozier N, Pallesen S, Aarø LE, et al. Sleep patterns and insomnia in young adults: a national survey of norwegian university students. *J Sleep Res* 2019 Apr;28(2):e12790 [FREE Full text] [doi: [10.1111/jsr.12790](https://doi.org/10.1111/jsr.12790)] [Medline: [30515935](https://pubmed.ncbi.nlm.nih.gov/30515935/)]

6. Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. *Sleep* 2013 Jul 01;36(7):1059-1068 [FREE Full text] [doi: [10.5665/sleep.2810](https://doi.org/10.5665/sleep.2810)] [Medline: [23814343](https://pubmed.ncbi.nlm.nih.gov/23814343/)]
7. Chiu H, Lee H, Chen P, Lai Y, Tu Y. Associations between sleep duration and suicidality in adolescents: a systematic review and dose-response meta-analysis. *Sleep Med Rev* 2018 Dec;42:119-126. [doi: [10.1016/j.smrv.2018.07.003](https://doi.org/10.1016/j.smrv.2018.07.003)] [Medline: [30093362](https://pubmed.ncbi.nlm.nih.gov/30093362/)]
8. Hysing M, Harvey AG, Linton SJ, Askeland KG, Sivertsen B. Sleep and academic performance in later adolescence: results from a large population-based study. *J Sleep Res* 2016 Jun;25(3):318-324 [FREE Full text] [doi: [10.1111/jsr.12373](https://doi.org/10.1111/jsr.12373)] [Medline: [26825591](https://pubmed.ncbi.nlm.nih.gov/26825591/)]
9. Wong MM, Brower KJ. The prospective relationship between sleep problems and suicidal behavior in the National Longitudinal Study of Adolescent Health. *J Psychiatr Res* 2012 Jul;46(7):953-959 [FREE Full text] [doi: [10.1016/j.jpsychires.2012.04.008](https://doi.org/10.1016/j.jpsychires.2012.04.008)] [Medline: [22551658](https://pubmed.ncbi.nlm.nih.gov/22551658/)]
10. Yang C, Wu C, Hsieh M, Liu M, Lu F. Coping with sleep disturbances among young adults: a survey of first-year college students in Taiwan. *Behav Med* 2003;29(3):133-138. [doi: [10.1080/08964280309596066](https://doi.org/10.1080/08964280309596066)] [Medline: [15206832](https://pubmed.ncbi.nlm.nih.gov/15206832/)]
11. Zhang J, Paksarian D, Lamers F, Hickie IB, He J, Merikangas KR. Sleep patterns and mental health correlates in US adolescents. *J Pediatr* 2017 Mar;182:137-143. [doi: [10.1016/j.jpeds.2016.11.007](https://doi.org/10.1016/j.jpeds.2016.11.007)] [Medline: [27939122](https://pubmed.ncbi.nlm.nih.gov/27939122/)]
12. Carskadon MA. Sleep in adolescents: the perfect storm. *Pediatr Clin North Am* 2011 Jun;58(3):637-647 [FREE Full text] [doi: [10.1016/j.pcl.2011.03.003](https://doi.org/10.1016/j.pcl.2011.03.003)] [Medline: [21600346](https://pubmed.ncbi.nlm.nih.gov/21600346/)]
13. Crowley SJ, Acebo C, Carskadon MA. Sleep, circadian rhythms, and delayed phase in adolescence. *Sleep Med* 2007 Sep;8(6):602-612. [doi: [10.1016/j.sleep.2006.12.002](https://doi.org/10.1016/j.sleep.2006.12.002)] [Medline: [17383934](https://pubmed.ncbi.nlm.nih.gov/17383934/)]
14. Gradisar M, Wolfson AR, Harvey AG, Hale L, Rosenberg R, Czeisler CA. The sleep and technology use of Americans: findings from the National Sleep Foundation's 2011 sleep in America poll. *J Clin Sleep Med* 2013 Dec 15;9(12):1291-1299 [FREE Full text] [doi: [10.5664/jcsm.3272](https://doi.org/10.5664/jcsm.3272)] [Medline: [24340291](https://pubmed.ncbi.nlm.nih.gov/24340291/)]
15. Randler C, Bilger S. Associations among sleep, chronotype, parental monitoring, and pubertal development among German adolescents. *J Psychol* 2009 Oct;143(5):509-520. [doi: [10.3200/JRL.143.5.509-520](https://doi.org/10.3200/JRL.143.5.509-520)] [Medline: [19943401](https://pubmed.ncbi.nlm.nih.gov/19943401/)]
16. Aoki H, Ozeki Y, Yamada N. Hypersensitivity of melatonin suppression in response to light in patients with delayed sleep phase syndrome. *Chronobiol Int* 2001 Mar;18(2):263-271. [doi: [10.1081/cbi-100103190](https://doi.org/10.1081/cbi-100103190)] [Medline: [11379666](https://pubmed.ncbi.nlm.nih.gov/11379666/)]
17. Carskadon MA, Vieira C, Acebo C. Association between puberty and delayed phase preference. *Sleep* 1993 Apr;16(3):258-262. [doi: [10.1093/sleep/16.3.258](https://doi.org/10.1093/sleep/16.3.258)] [Medline: [8506460](https://pubmed.ncbi.nlm.nih.gov/8506460/)]
18. James JE, Kristjánsson AL, Sigfúsdóttir ID. Adolescent substance use, sleep, and academic achievement: evidence of harm due to caffeine. *J Adolesc* 2011 Aug;34(4):665-673. [doi: [10.1016/j.adolescence.2010.09.006](https://doi.org/10.1016/j.adolescence.2010.09.006)] [Medline: [20970177](https://pubmed.ncbi.nlm.nih.gov/20970177/)]
19. Gibson ES, Powles AC, Thabane L, O'Brien S, Molnar DS, Trajanovic N, et al. "Sleepiness" is serious in adolescence: two surveys of 3235 Canadian students. *BMC Public Health* 2006 May 02;6:116 [FREE Full text] [doi: [10.1186/1471-2458-6-116](https://doi.org/10.1186/1471-2458-6-116)] [Medline: [16670019](https://pubmed.ncbi.nlm.nih.gov/16670019/)]
20. Pagel JF, Forister N, Kwiatkowi C. Adolescent sleep disturbance and school performance: the confounding variable of socioeconomic status. *J Clin Sleep Med* 2007 Feb 15;3(1):19-23. [Medline: [17557448](https://pubmed.ncbi.nlm.nih.gov/17557448/)]
21. Qaseem A, Kansagara D, Forcica MA, Cooke M, Denberg TD, Clinical Guidelines Committee of the American College of Physicians. Management of chronic insomnia disorder in adults: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2016 Jul 19;165(2):125-133 [FREE Full text] [doi: [10.7326/M15-2175](https://doi.org/10.7326/M15-2175)] [Medline: [27136449](https://pubmed.ncbi.nlm.nih.gov/27136449/)]
22. Insomnia. National Institute for HealthCare Excellence (NICE). 2015. URL: <https://cks.nice.org.uk/insomnia#!management> [accessed 2020-10-20]
23. Owens JA, Rosen CL, Mindell JA. Medication use in the treatment of pediatric insomnia: results of a survey of community-based pediatricians. *Pediatrics* 2003 May;111(5 Pt 1):628-635. [doi: [10.1542/peds.111.5.e628](https://doi.org/10.1542/peds.111.5.e628)] [Medline: [12728122](https://pubmed.ncbi.nlm.nih.gov/12728122/)]
24. Mindell JA, Kuhn B, Lewin DS, Meltzer LJ, Sadeh A, American Academy of Sleep Medicine. Behavioral treatment of bedtime problems and night wakings in infants and young children. *Sleep* 2006 Oct;29(10):1263-1276. [doi: [10.1093/SLEEP/29.10.1263](https://doi.org/10.1093/SLEEP/29.10.1263)] [Medline: [17068979](https://pubmed.ncbi.nlm.nih.gov/17068979/)]
25. Stojanovski SD, Rasu RS, Balkrishnan R, Nahata MC. Trends in medication prescribing for pediatric sleep difficulties in US outpatient settings. *Sleep* 2007 Aug;30(8):1013-1017 [FREE Full text] [doi: [10.1093/sleep/30.8.1013](https://doi.org/10.1093/sleep/30.8.1013)] [Medline: [17702271](https://pubmed.ncbi.nlm.nih.gov/17702271/)]
26. Kloss JD, Nash CO, Horsey SE, Taylor DJ. The delivery of behavioral sleep medicine to college students. *J Adolesc Health* 2011 Jun;48(6):553-561. [doi: [10.1016/j.jadohealth.2010.09.023](https://doi.org/10.1016/j.jadohealth.2010.09.023)] [Medline: [21575813](https://pubmed.ncbi.nlm.nih.gov/21575813/)]
27. Harvey AG, Hein K, Dolsen MR, Dong L, Rabe-Hesketh S, Gumport NB, et al. Modifying the impact of eveningness chronotype ("Night-owls") in youth: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry* 2018 Oct;57(10):742-754. [doi: [10.1016/j.jaac.2018.04.020](https://doi.org/10.1016/j.jaac.2018.04.020)] [Medline: [30274649](https://pubmed.ncbi.nlm.nih.gov/30274649/)]
28. Gumport NB, Dolsen MR, Harvey AG. Usefulness and utilization of treatment elements from the Transdiagnostic Sleep and Circadian Intervention for adolescents with an evening circadian preference. *Behav Res Ther* 2019 Dec;123:103504. [doi: [10.1016/j.brat.2019.103504](https://doi.org/10.1016/j.brat.2019.103504)] [Medline: [31678861](https://pubmed.ncbi.nlm.nih.gov/31678861/)]

29. Dong L, Gumpert NB, Martinez AJ, Harvey AG. Is improving sleep and circadian problems in adolescence a pathway to improved health? A mediation analysis. *J Consult Clin Psychol* 2019 Sep;87(9):757-771 [FREE Full text] [doi: [10.1037/ccp0000423](https://doi.org/10.1037/ccp0000423)] [Medline: [31246052](https://pubmed.ncbi.nlm.nih.gov/31246052/)]
30. Koffel E, Amundson E, Polusny G, Wisdom JP. "You're missing out on something great": patient and provider perspectives on increasing the use of cognitive behavioral therapy for insomnia. *Behav Sleep Med* 2020;18(3):358-371. [doi: [10.1080/15402002.2019.1591958](https://doi.org/10.1080/15402002.2019.1591958)] [Medline: [30907144](https://pubmed.ncbi.nlm.nih.gov/30907144/)]
31. Boerner KE, Coulombe JA, Corkum P. Barriers and facilitators of evidence-based practice in pediatric behavioral sleep care: qualitative analysis of the perspectives of health professionals. *Behav Sleep Med* 2015;13(1):36-51. [doi: [10.1080/15402002.2013.838766](https://doi.org/10.1080/15402002.2013.838766)] [Medline: [24364693](https://pubmed.ncbi.nlm.nih.gov/24364693/)]
32. Harvey AG, Buysse DJ. *Treating Sleep Problems: A Transdiagnostic Approach*. 1st Edition. New York, NY: The Guilford Press; 2017.
33. Crookston BT, West JH, Hall PC, Dahle KM, Heaton TL, Beck RN, et al. Mental and emotional self-help technology apps: cross-sectional study of theory, technology, and mental health behaviors. *JMIR Ment Health* 2017 Oct 17;4(4):e45 [FREE Full text] [doi: [10.2196/mental.7262](https://doi.org/10.2196/mental.7262)] [Medline: [29042340](https://pubmed.ncbi.nlm.nih.gov/29042340/)]
34. Radovic A, McCarty CA, Katzman K, Richardson LP. Adolescents' perspectives on using technology for health: qualitative study. *JMIR Pediatr Parent* 2018;1(1):e2 [FREE Full text] [doi: [10.2196/pediatrics.8677](https://doi.org/10.2196/pediatrics.8677)] [Medline: [30740590](https://pubmed.ncbi.nlm.nih.gov/30740590/)]
35. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: a survey of youths' web-based resource preferences in Canada. *J Med Internet Res* 2014 Dec 17;16(12):e293 [FREE Full text] [doi: [10.2196/jmir.3526](https://doi.org/10.2196/jmir.3526)] [Medline: [25519847](https://pubmed.ncbi.nlm.nih.gov/25519847/)]
36. Mindell JA, Du Mond CE, Sadeh A, Telofski LS, Kulkarni N, Gunn E. Efficacy of an internet-based intervention for infant and toddler sleep disturbances. *Sleep* 2011 Apr 01;34(4):451-458 [FREE Full text] [doi: [10.1093/sleep/34.4.451](https://doi.org/10.1093/sleep/34.4.451)] [Medline: [21461323](https://pubmed.ncbi.nlm.nih.gov/21461323/)]
37. Corkum PV, Reid GJ, Hall WA, Godbout R, Stremler R, Weiss SK, et al. Evaluation of an internet-based behavioral intervention to improve psychosocial health outcomes in children with insomnia (Better Nights, Better Days): protocol for a randomized controlled trial. *JMIR Res Protoc* 2018 Mar 26;7(3):e76 [FREE Full text] [doi: [10.2196/resprot.8348](https://doi.org/10.2196/resprot.8348)] [Medline: [29581089](https://pubmed.ncbi.nlm.nih.gov/29581089/)]
38. Speth TA, Coulombe JA, Markovich AN, Chambers CT, Godbout R, Gruber R, et al. Barriers, facilitators, and usability of an internet intervention for children aged 1 to 10 years with insomnia. *Transl Issues Psychol Sci* 2015;1(1):16-31. [doi: [10.1037/tps0000016](https://doi.org/10.1037/tps0000016)]
39. Werner-Seidler A, Wong Q, Johnston L, O'Dea B, Torok M, Christensen H. Pilot evaluation of the Sleep Ninja: a smartphone application for adolescent insomnia symptoms. *BMJ Open* 2019 May 27;9(5):e026502 [FREE Full text] [doi: [10.1136/bmjopen-2018-026502](https://doi.org/10.1136/bmjopen-2018-026502)] [Medline: [31133584](https://pubmed.ncbi.nlm.nih.gov/31133584/)]
40. Werner-Seidler A, O'Dea B, Shand F, Johnston L, Frayne A, Fogarty AS, et al. A smartphone app for adolescents with sleep disturbance: development of the Sleep Ninja. *JMIR Ment Health* 2017 Jul 28;4(3):e28 [FREE Full text] [doi: [10.2196/mental.7614](https://doi.org/10.2196/mental.7614)] [Medline: [28754651](https://pubmed.ncbi.nlm.nih.gov/28754651/)]
41. de Bruin EJ, Bögels SM, Oort FJ, Meijer AM. Efficacy of cognitive behavioral therapy for insomnia in adolescents: a randomized controlled trial with internet therapy, group therapy and a waiting list condition. *Sleep* 2015 Dec 01;38(12):1913-1926 [FREE Full text] [doi: [10.5665/sleep.5240](https://doi.org/10.5665/sleep.5240)] [Medline: [26158889](https://pubmed.ncbi.nlm.nih.gov/26158889/)]
42. Espie CA, Kyle SD, Williams C, Ong JC, Douglas NJ, Hames P, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. *Sleep* 2012 Jun 01;35(6):769-781 [FREE Full text] [doi: [10.5665/sleep.1872](https://doi.org/10.5665/sleep.1872)] [Medline: [22654196](https://pubmed.ncbi.nlm.nih.gov/22654196/)]
43. DOZE - delivering online z's. DOZE App. 2019. URL: <https://dozeapp.ca/> [accessed 2020-10-20]
44. Hackett CL, Mulvale G, Miatello A. Co-designing for quality: creating a user-driven tool to improve quality in youth mental health services. *Health Expect* 2018 Dec;21(6):1013-1023 [FREE Full text] [doi: [10.1111/hex.12694](https://doi.org/10.1111/hex.12694)] [Medline: [29707865](https://pubmed.ncbi.nlm.nih.gov/29707865/)]
45. Harrington C, Wilcox L, Connelly K, Rogers W, Sanford J. Designing health and fitness apps with older adults: examining the value of experience-based co-design. In: *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare.: Association for Computing Machinery; 2018 Presented at: PervasiveHealth '18: 12th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 21 - 24, 2018; New York, USA p. 15-24.* [doi: [10.1145/3240925.3240929](https://doi.org/10.1145/3240925.3240929)]
46. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of eHealth interventions With children and young people. *Front Psychiatry* 2018;9:481 [FREE Full text] [doi: [10.3389/fpsy.2018.00481](https://doi.org/10.3389/fpsy.2018.00481)] [Medline: [30405450](https://pubmed.ncbi.nlm.nih.gov/30405450/)]
47. Shah LM, Yang WE, Demo RC, Lee MA, Weng D, Shan R, et al. Technical guidance for clinicians interested in partnering with engineers in mobile health development and evaluation. *JMIR Mhealth Uhealth* 2019 May 15;7(5):e14124 [FREE Full text] [doi: [10.2196/14124](https://doi.org/10.2196/14124)] [Medline: [31094337](https://pubmed.ncbi.nlm.nih.gov/31094337/)]
48. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep* 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: [10.5665/sleep.1642](https://doi.org/10.5665/sleep.1642)] [Medline: [22294820](https://pubmed.ncbi.nlm.nih.gov/22294820/)]
49. Borkovec TD, Nau SD. Credibility of analogue therapy rationales. *J Behav Ther Exp Psychiatry* 1972 Dec;3(4):257-260. [doi: [10.1016/0005-7916\(72\)90045-6](https://doi.org/10.1016/0005-7916(72)90045-6)]

50. Carney CE, Edinger JD, Kuchibhatla M, Lachowski AM, Bogouslavsky O, Krystal AD, et al. Cognitive behavioral insomnia therapy for those with insomnia and depression: a randomized controlled clinical trial. *Sleep* 2017 Apr 01;40(4):zsx019 [FREE Full text] [doi: [10.1093/sleep/zsx019](https://doi.org/10.1093/sleep/zsx019)] [Medline: [28199710](https://pubmed.ncbi.nlm.nih.gov/28199710/)]
51. Morin C. *Insomnia: Psychological Assessment and Management*. New York, NY: Guilford Press; 1993.
52. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep* 2011 May 01;34(5):601-608. [doi: [10.1093/sleep/34.5.601](https://doi.org/10.1093/sleep/34.5.601)] [Medline: [21532953](https://pubmed.ncbi.nlm.nih.gov/21532953/)]
53. Chung KF, Kan KK, Yeung WF. Assessing insomnia in adolescents: comparison of Insomnia Severity Index, Athens Insomnia Scale and Sleep Quality Index. *Sleep Med* 2011 May;12(5):463-470. [doi: [10.1016/j.sleep.2010.09.019](https://doi.org/10.1016/j.sleep.2010.09.019)] [Medline: [21493134](https://pubmed.ncbi.nlm.nih.gov/21493134/)]
54. Chahoud M, Chahine R, Salameh P, Sauleau EA. Reliability, factor analysis and internal consistency calculation of the Insomnia Severity Index (ISI) in French and in English among Lebanese adolescents. *eNeurologicalSci* 2017 Mar 18;7:9-14 [FREE Full text] [doi: [10.1016/j.ensci.2017.03.003](https://doi.org/10.1016/j.ensci.2017.03.003)] [Medline: [29260018](https://pubmed.ncbi.nlm.nih.gov/29260018/)]
55. Gerber M, Lang C, Lemola S, Colledge F, Kalak N, Holsboer-Trachsler E, et al. Validation of the German version of the Insomnia Severity Index in adolescents, young adults and adult workers: results from three cross-sectional studies. *BMC Psychiatry* 2016 May 31;16:174 [FREE Full text] [doi: [10.1186/s12888-016-0876-8](https://doi.org/10.1186/s12888-016-0876-8)] [Medline: [27245844](https://pubmed.ncbi.nlm.nih.gov/27245844/)]
56. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001 Jul;2(4):297-307. [doi: [10.1016/s1389-9457\(00\)00065-4](https://doi.org/10.1016/s1389-9457(00)00065-4)] [Medline: [11438246](https://pubmed.ncbi.nlm.nih.gov/11438246/)]
57. Smith CS, Reilly C, Midkiff K. Evaluation of three circadian rhythm questionnaires with suggestions for an improved measure of morningness. *J Appl Psychol* 1989 Oct;74(5):728-738. [doi: [10.1037/0021-9010.74.5.728](https://doi.org/10.1037/0021-9010.74.5.728)] [Medline: [2793773](https://pubmed.ncbi.nlm.nih.gov/2793773/)]
58. Smith CS, Folkard S, Schmieder RA, Parra LF, Spelten E, Almiral H, et al. Investigation of morning-evening orientation in six countries using the preferences scale. *Pers Individ Dif* 2002 Apr;32(6):949-968. [doi: [10.1016/s0191-8869\(01\)00098-8](https://doi.org/10.1016/s0191-8869(01)00098-8)]
59. Randler C. Morningness-eveningness comparison in adolescents from different countries around the world. *Chronobiol Int* 2008 Nov;25(6):1017-1028. [doi: [10.1080/07420520802551519](https://doi.org/10.1080/07420520802551519)] [Medline: [19005902](https://pubmed.ncbi.nlm.nih.gov/19005902/)]
60. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989 Oct;46(10):1121-1123. [doi: [10.1001/archneur.1989.00520460115022](https://doi.org/10.1001/archneur.1989.00520460115022)] [Medline: [2803071](https://pubmed.ncbi.nlm.nih.gov/2803071/)]
61. Huang Y, Katz BZ, Mears C, Kielhofner GW, Taylor R. Postinfectious fatigue in adolescents and physical activity. *Arch Pediatr Adolesc Med* 2010 Sep;164(9):803-809 [FREE Full text] [doi: [10.1001/archpediatrics.2010.144](https://doi.org/10.1001/archpediatrics.2010.144)] [Medline: [20819961](https://pubmed.ncbi.nlm.nih.gov/20819961/)]
62. Gordon BA, Knapman LM, Lubitz L. Graduated exercise training and progressive resistance training in adolescents with chronic fatigue syndrome: a randomized controlled pilot study. *Clin Rehabil* 2010 Dec;24(12):1072-1079. [doi: [10.1177/0269215510371429](https://doi.org/10.1177/0269215510371429)] [Medline: [20605858](https://pubmed.ncbi.nlm.nih.gov/20605858/)]
63. Spilsbury JC, Drotar D, Rosen CL, Redline S. The Cleveland Adolescent Sleepiness Questionnaire: a new measure to assess excessive daytime sleepiness in adolescents. *J Clin Sleep Med* 2007 Oct 15;3(6):603-612 [FREE Full text] [doi: [10.5664/jcsm.26971](https://doi.org/10.5664/jcsm.26971)] [Medline: [17993042](https://pubmed.ncbi.nlm.nih.gov/17993042/)]
64. Haroz EE, Ybarra ML, Eaton WW. Psychometric evaluation of a self-report scale to measure adolescent depression: the CESDR-10 in two national adolescent samples in the United States. *J Affect Disord* 2014 Apr;158:154-160 [FREE Full text] [doi: [10.1016/j.jad.2014.02.009](https://doi.org/10.1016/j.jad.2014.02.009)] [Medline: [24655780](https://pubmed.ncbi.nlm.nih.gov/24655780/)]
65. Ree M, MacLeod C, French D, Locke V. The State-Trait Inventory for Cognitive and Somatic Anxiety: development and validation. *State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA) 2000:A*. [doi: [10.1037/t01184-000](https://doi.org/10.1037/t01184-000)]
66. Deacy AD, Gayes LA, De Lurgio S, Wallace DP. Adaptation of the State-Trait Inventory for Cognitive and Somatic Anxiety for use in children: a preliminary analysis. *J Pediatr Psychol* 2016 Oct;41(9):1033-1043. [doi: [10.1093/jpepsy/jsw030](https://doi.org/10.1093/jpepsy/jsw030)] [Medline: [27114070](https://pubmed.ncbi.nlm.nih.gov/27114070/)]
67. Grös DF, Antony MM, Simms LJ, McCabe RE. Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): comparison to the State-Trait Anxiety Inventory (STAI). *Psychol Assess* 2007 Dec;19(4):369-381. [doi: [10.1037/1040-3590.19.4.369](https://doi.org/10.1037/1040-3590.19.4.369)] [Medline: [18085930](https://pubmed.ncbi.nlm.nih.gov/18085930/)]
68. Roberts K. *State-Trait Inventory for Cognitive and Somatic Anxiety: psychometric properties and experimental manipulation to evaluate sensitivity to change and predictive validity*. Thesis and Dissertations - York University, Toronto. 2013. URL: https://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/31982/Roberts_Karen_E_2013_PhD.pdf?isAllowed=y&sequence=1 [accessed 2020-10-20]
69. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473-483. [doi: [10.1097/00005650-199206000-00002](https://doi.org/10.1097/00005650-199206000-00002)] [Medline: [1593914](https://pubmed.ncbi.nlm.nih.gov/1593914/)]
70. McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994 Jan;32(1):40-66. [doi: [10.1097/00005650-199401000-00004](https://doi.org/10.1097/00005650-199401000-00004)] [Medline: [8277801](https://pubmed.ncbi.nlm.nih.gov/8277801/)]
71. Taylor DJ, Bramoweth AD, Grieser EA, Tatum JI, Roane BM. Epidemiology of insomnia in college students: relationship with mental health, quality of life, and substance use difficulties. *Behav Ther* 2013 Sep;44(3):339-348. [doi: [10.1016/j.beth.2012.12.001](https://doi.org/10.1016/j.beth.2012.12.001)] [Medline: [23768662](https://pubmed.ncbi.nlm.nih.gov/23768662/)]
72. Benbassat J, Pilpel D, Tidhar M. Patients' preferences for participation in clinical decision making: a review of published surveys. *Behav Med* 1998;24(2):81-88. [doi: [10.1080/08964289809596384](https://doi.org/10.1080/08964289809596384)] [Medline: [9695899](https://pubmed.ncbi.nlm.nih.gov/9695899/)]

73. 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/)]
74. Gowin M, Cheney M, Gwin S, Franklin Wann T. Health and fitness app use in college students: a qualitative study. *Am J Health Educ* 2015 Jul 06;46(4):223-230. [doi: [10.1080/19325037.2015.1044140](https://doi.org/10.1080/19325037.2015.1044140)]
75. Miller KE, Kuhn E, Owen JE, Taylor K, Yu JS, Weiss BJ, et al. Clinician perceptions related to the use of the CBT-I Coach mobile app. *Behav Sleep Med* 2019;17(4):481-491. [doi: [10.1080/15402002.2017.1403326](https://doi.org/10.1080/15402002.2017.1403326)] [Medline: [29120247](https://pubmed.ncbi.nlm.nih.gov/29120247/)]
76. Cliffe B, Croker A, Denne M, Smith J, Stallard P. Digital cognitive behavioral therapy for insomnia for adolescents with mental health problems: feasibility open trial. *JMIR Ment Health* 2020 Mar 03;7(3):e14842 [FREE Full text] [doi: [10.2196/14842](https://doi.org/10.2196/14842)] [Medline: [32134720](https://pubmed.ncbi.nlm.nih.gov/32134720/)]
77. Voerman JS, Remerie S, Westendorp T, Timman R, Busschbach JJ, Passchier J, et al. Effects of a guided internet-delivered self-help intervention for adolescents with chronic pain. *J Pain* 2015 Nov;16(11):1115-1126. [doi: [10.1016/j.jpain.2015.07.011](https://doi.org/10.1016/j.jpain.2015.07.011)] [Medline: [26281947](https://pubmed.ncbi.nlm.nih.gov/26281947/)]
78. Koffel E, Kuhn E, Petsoulis N, Erbes CR, Anders S, Hoffman JE, et al. A randomized controlled pilot study of CBT-I Coach: feasibility, acceptability, and potential impact of a mobile phone application for patients in cognitive behavioral therapy for insomnia. *Health Informatics J* 2018 Mar;24(1):3-13. [doi: [10.1177/1460458216656472](https://doi.org/10.1177/1460458216656472)] [Medline: [27354394](https://pubmed.ncbi.nlm.nih.gov/27354394/)]

Abbreviations

AYA: adolescent and young adult

CASQ: Cleveland Adolescent Sleepiness Questionnaire

CBT-I: cognitive behavioral therapy for insomnia

CESDR-10: Center for Epidemiologic Studies Depression Scale–Revised 10-item Version for Adolescents

CSM: Composite Scale of Morningness

DOZE: Delivering Online Zzz's with Empirical Support

FSS: Fatigue Severity Scale

ISI: Insomnia Severity Index

MD: mean difference

SE: sleep efficiency

SF-36: RAND 36-item Short Form Health Survey 1.0

SOL: sleep onset latency

STICSA: State-Trait Inventory of Cognitive and Somatic Anxiety

TEQ: Treatment Evaluation Questionnaire

TIB: time in bed

TST: total sleep time

TWT: total wake time

Edited by G Eysenbach; submitted 30.10.20; peer-reviewed by J Yu, D Dunsmuir, B Cliffe; comments to author 18.11.20; revised version received 06.01.21; accepted 13.09.21; published 01.11.21.

Please cite as:

Carmona NE, Usyatynsky A, Kutana S, Corkum P, Henderson J, McShane K, Shapiro C, Sidani S, Stinson J, Carney CE
A Transdiagnostic Self-management Web-Based App for Sleep Disturbance in Adolescents and Young Adults: Feasibility and Acceptability Study

JMIR Form Res 2021;5(11):e25392

URL: <https://formative.jmir.org/2021/11/e25392>

doi: [10.2196/25392](https://doi.org/10.2196/25392)

PMID: [34723820](https://pubmed.ncbi.nlm.nih.gov/34723820/)

©Nicole E Carmona, Aleksandra Usyatynsky, Samlau Kutana, Penny Corkum, Joanna Henderson, Kelly McShane, Colin Shapiro, Souraya Sidani, Jennifer Stinson, Colleen E Carney. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 01.11.2021. 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

Organizational Readiness for Implementing an Internet-Based Cognitive Behavioral Therapy Intervention for Depression Across Community Mental Health Services in Albania and Kosovo: Directed Qualitative Content Analysis

Asmae Doukani^{1*}, BSc, MSc; Arlinda Cerga Pashoja^{1,2*}, BSc, MSc, PhD; Naim Fanaj³, BSc, MSc, PhD; Gentiana Qirjako^{4,5,6}, BSc, MSc, PhD; Andia Meksi^{4,5}, BSc, MSc; Sevim Mustafa³, BSc, MSc, PhD; Christiaan Vis^{7,8,9}, BSc, MSc; Juliane Hug¹⁰, BSc, MSc

¹Department of Population Health, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, United Kingdom

²Global Public Health, United Kingdom Health Security Agency, London, United Kingdom

³Mental Health Centre, Prizren, Kosovo

⁴Community Centre for Health and Wellbeing, Tirana, Albania

⁵Department of Promotion, Institute of Public Health, Tirana, Albania

⁶Department of Public Health, Faculty of Medicine, University of Medicine, Tirana, Albania

⁷Department of Clinical, Neuro & Developmental Psychology, Faculty of Behavioural and Movement Sciences, VU Amsterdam, Amsterdam, Netherlands

⁸Mental Health, Amsterdam Public Health Research Institute, Amsterdam, Netherlands

⁹World Health Organization, Collaborating Centre for Research and Dissemination of Psychological Interventions, Geneva, Switzerland

¹⁰European Alliance Against Depression, Leipzig, Germany

*these authors contributed equally

Corresponding Author:

Arlinda Cerga Pashoja, BSc, MSc, PhD

Department of Population Health

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Keppel Street

London, WC1E 7HT

United Kingdom

Phone: 44 7787447507

Email: arlinda.cerga-pashoja@lshtm.ac.uk

Abstract

Background: The use of digital mental health programs such as internet-based cognitive behavioral therapy (iCBT) holds promise in increasing the quality and access of mental health services. However very little research has been conducted in understanding the feasibility of implementing iCBT in Eastern Europe.

Objective: The aim of this study was to qualitatively assess organizational readiness for implementing iCBT for depression within community mental health centers (CMHCs) across Albania and Kosovo.

Methods: We used qualitative semistructured focus group discussions that were guided by Bryan Weiner's model of organizational readiness for implementing change. The questions broadly explored shared determination to implement change (change commitment) and shared belief in their collective capability to do so (change efficacy). Data were collected between November and December 2017. A range of health care professionals working in and in association with CMHCs were recruited from 3 CMHCs in Albania and 4 CMHCs in Kosovo, which were participating in a large multinational trial on the implementation of iCBT across 9 countries (Horizon 2020 ImpleMentAll project). Data were analyzed using a directed approach to qualitative content analysis, which used a combination of both inductive and deductive approaches.

Results: Six focus group discussions involving 69 mental health care professionals were conducted. Participants from Kosovo (36/69, 52%) and Albania (33/69, 48%) were mostly females (48/69, 70%) and nurses (26/69, 38%), with an average age of 41.3

years. A directed qualitative content analysis revealed several barriers and facilitators potentially affecting the implementation of digital CBT interventions for depression in community mental health settings. While commitment for change was high, change efficacy was limited owing to a range of situational factors. Barriers impacting “change efficacy” included lack of clinical fit for iCBT, high stigma affecting help-seeking behaviors, lack of human resources, poor technological infrastructure, and high caseload. Facilitators included having a high interest and capability in receiving training for iCBT. For “change commitment,” participants largely expressed welcoming innovation and that iCBT could increase access to treatments for geographically isolated people and reduce the stigma associated with mental health care.

Conclusions: In summary, participants perceived iCBT positively in relation to promoting innovation in mental health care, increasing access to services, and reducing stigma. However, a range of barriers was also highlighted in relation to accessing the target treatment population, a culture of mental health stigma, underdeveloped information and communications technology infrastructure, and limited appropriately trained health care workforce, which reduce organizational readiness for implementing iCBT for depression. Such barriers may be addressed through (1) a public-facing campaign that addresses mental health stigma, (2) service-level adjustments that permit staff with the time, resources, and clinical supervision to deliver iCBT, and (3) establishment of a suitable clinical training curriculum for health care professionals.

Trial Registration: ClinicalTrials.gov NCT03652883; <https://clinicaltrials.gov/ct2/show/NCT03652883>

(*JMIR Form Res* 2021;5(11):e29280) doi:[10.2196/29280](https://doi.org/10.2196/29280)

KEYWORDS

e-mental health; digital mental health; internet-based cognitive behavioral therapy; implementation science; organizational readiness for implementing change; community mental health center; qualitative interviews; content analysis; Albania and Kosovo

Introduction

Albania and Kosovo are upper-middle-income countries in the Southeast of Europe. Situated in close proximity to high-income countries with well-resourced health care systems, the state of mental health care in Eastern Europe has gone unnoticed and has been referred to as a global blind spot [1]. In 2017, the burden of mental illness in Albania was estimated to be 3888 per 100,000 people, with disproportionately lower human resource availability, including only 1 psychiatrist, 1 psychologist, and 7 nurses per 100,000 people [2]. In Kosovo, the rate of mental illness has been notably higher due to the conflict that took place in 1998-1999 [3,4]. A survey conducted by the Kosovo Rehabilitation Centre for Torture Victims found that 27.7% of the population reported substantial psychiatric morbidity, in which 22% presented with symptoms of posttraumatic stress disorder, 41.8% for depression, and 43.1% for emotional distress [3]. Another study found that 64.9% of the Kosovan population reported having traumatic experiences during the war, resulting in 200,000-400,000 traumatized persons, which was additional to the existing figure of persons with mental illness [4]. The prevalence of mental illness in postwar Kosovo continues to be a problem, with only a slow decline to show years on after the conflict [5-7]. Despite the high burden of mental illness, the mental health workforce in Kosovo is comparatively lower than that in the neighboring countries, while the country's mental health budget only equates to 2% of the average mental health budgets of countries in the European region [2].

Limited resources for mental health is, however, not the only barrier toward addressing the care gap. A review of mental health care systems in eastern Europe found that Central and Eastern Europe experienced higher reports of public stigma associated with mental illness compared to other European countries [8]. The high level of public stigma (unwillingness to accept people with severe mental illnesses) may therefore have

far reaching implications that can negatively influence policy, funding, recovery, help-seeking behaviors, service quality, and quality of life for people with mental health conditions [8]. A study on the factors that influence access to mental health services in Southeastern Europe (Romania, Bulgaria, and Albania) found that mental health stigma and a lack of knowledge around mental illness were among the factors that influenced delayed access to mental health services [9].

As a response to mental health care access barriers, the use of digital technologies has been put forward as a viable approach for closing the care gap in mental health care owing to the greater potential of technology to expand the mental health workforce, increase access to evidence-based interventions at a lower cost, reduce stigma due to opportunities to engage in treatment remotely, and enable interventions to be linguistically and culturally adapted [10]. Internet penetration rates for Albania (72%) and Kosovo (89%) [11,12] support the argument for harnessing digital technologies in mental health care, as rates appear to be close to or exceed the average internet penetration in high-income countries (75%). However, little to nothing is known about the readiness of mental health services in implementing digital mental health care interventions. Readiness or organizational readiness can be defined as “the extent to which organizational members are psychologically and behaviorally prepared to implement organizational change” [13]. The concept of organizational readiness is an integral component of implementing new health programs owing to the growing recognition that programs may fail not as a result of the digital health innovation but because organizational readiness for change is not adequately evaluated and addressed [14,15]. The assessment of organizational readiness is therefore advised in the early phases of implementation in order to gain a better understanding of the challenges and facilitators for successful implementation, augmentation, and optimization of implementation strategies [16].

The aim of this study was to undertake a qualitative examination of organizational readiness for implementing an internet-based cognitive behavioral therapy (iCBT) intervention for people with depression in 7 community mental health centers (CMHCs) across Albania and Kosovo who were partaking in a large multinational trial on the implementation of iCBT (ImpleMentAll trial, further information about the trial can be found in the Methods section) [17]. To our knowledge, this study will be the first to explore organizational readiness for implementing iCBT as well as a mental health service more broadly in Albania and Kosovo.

Methods

Participants and Procedure

Qualitative focus group discussions (FGDs) were conducted with mental health care stakeholders involved on the ImpleMentAll study, a multinational trial funded by the European Union's Horizon 2020 program. The project aimed to (1) develop and apply tailored implementation strategies for implementing evidence-based iCBT services for common mental disorders in routine mental health care through and (2) conduct a stepped wedge-cluster randomized trial to investigate the effectiveness of tailored implementation, when compared to implementation as usual (more information about the trial can be found in the trial protocol by Bührmann and colleagues [17]). Participants were recruited from 7 CMHCs across Albania (located in Tirana, Shkoder, and Korce) and Kosovo (Prizren, Gjilan, Prishtine, and Mitrovice). Focus groups were conducted prior to the commencement of the ImpleMentAll trial. A purposive sampling method was used for data collection to facilitate access to key informants (mental health care professionals working in and in association with CMHC) and maximum variations within an organization [18,19] considering diversity across age, gender, job role, and the level of experience of working with people with depression [19]. FGDs were conducted in a meeting room within respective CMHCs.

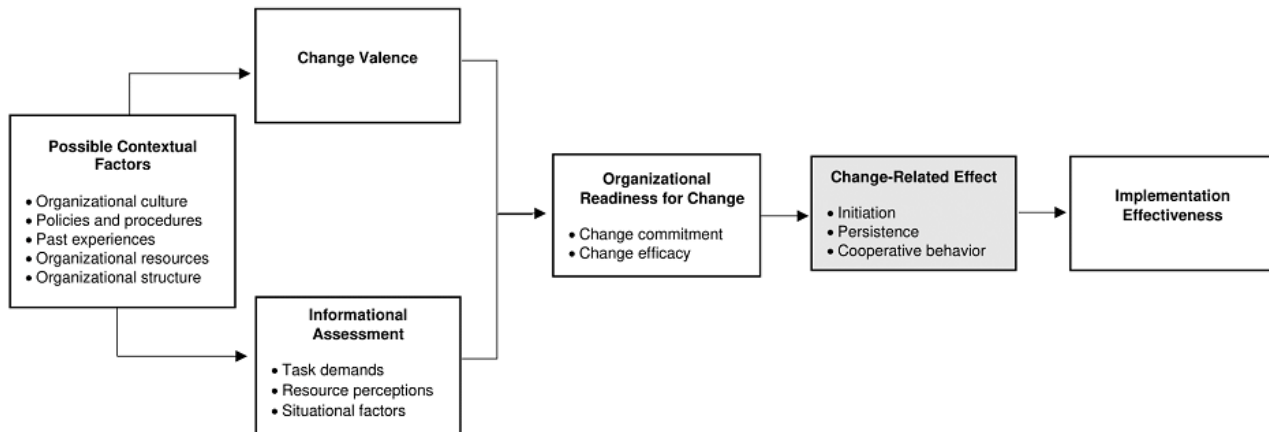
FGD topic guides were loosely directed by Weiner's [13] concept of organizational readiness for implementing change (ORIC) [20]. For example, we asked "What kind of training do staff that provide treatment for depression have?" which aims to understand specific organizational resources that are important for the effective implementation of iCBT in services [13]. FGD interviews were facilitated by ACP, GQ, and AM in Albania and ACP, NF, and SM in Kosovo. In Albania, interviews were audio recorded with a digital voice recorder. In Kosovo, the participants requested for discussion not to be audio recorded; therefore, interviews were captured in writing and special efforts were made to document the interviews verbatim in real time.

Audio recordings of the interviews were then transcribed verbatim in Albanian and subsequently translated into English. The notes made in Kosovan were also translated to English. This study was approved by the Republic of Albania Ministry of Health Social Protection Ethics Committee on November 17, 2017, and the Republic of Kosovo Qeveria-Vlada Government on February 11, 2017. This project was funded by the European Union's Horizon 2020 research and innovation program under grant agreement 733025 and received funding from the National Health and Medical Research Council European Union program by the Australian Government (1142363). Funding bodies had no influence on the design of this study. The trial registration number is ClinicalTrials.gov NCT03652883.

Analytical Framework

ORIC, as theorized by Bryan Weiner [13], was used as an analytical framework, as outlined in Figure 1. ORIC [13,21] is a multilevel and multifaceted construct that refers to the organizational members' shared determination to implement a change (change commitment) and shared belief in their collective capability to do so (change efficacy). Readiness can be theorized, assessed, and studied at the unit of the individual, group, department, and at the organization (eg, CMHC) level. ORIC varies as a function of how much organizational members value the change and how favorably they appraise 3 key determinants of implementation capability: (1) task demands, (2) resource availability, and (3) situational factors. High levels of organizational readiness for change is indicated by organizational members who are more likely to initiate change, exert greater efforts and persistence, and display more cooperative behaviors. The categories outlined in Weiner's [13] theory were reviewed by the research team to ensure that they align with the preimplementation context that the FGDs took place in. A decision was made to exclude all categories under the change-related effect (grey box in Figure 1), as efforts to initiate change were not experienced at the time in which the FGDs were conducted. Organizational readiness can be assessed across different units, including the individual (individual-level) or the team, department, or organization (supraindividual level). Considering that improvements in health care delivery often involve changes in collective behavior (eg, in staffing, working processes and procedures, decision making, communication), our study will primarily focus on assessing organizational readiness for change at a supraindividual level in relation to the organization across multiple CMHCs that while serving similar clinical populations also vary in context. However, given that opinions may vary between organizations and individuals, our study will also assess organizational readiness for change at an individual level [13,21,22].

Figure 1. Analytical framework on the determinants and outcomes of organizational readiness for change developed by Bryan Weiner [13].



Data Analysis

Data were analyzed at 2 levels: (1) manifest content analysis, which refers to the use of only transcribed interview text, and (2) latent content analysis, which relies on the reflections and interpretations [23]. Thomas and Magilvy [24] suggest that using both types of analyses is important for developing a deep understanding of the data. The unit of analysis for assessing ORIC within the qualitative interview was the CMHC [25]. Prior to data analysis, qualitative data analysts were provided with a sufficient scientific and content-based mastery of the ORIC theory [13], which was used to heuristically guide the data analysis. Definitions of ORIC categories are outlined in Figure 1. The directed content analysis commenced with immersion in the data by all authors of this paper. In order to enhance reliability, the analysis and coding were carried out independently and simultaneously in Albanian (ACP, GQ, AM, NF, and SM) and in English by AD as well as researchers from the ImpleMentAll trial who were not directly involved in the study (CV and JH) [26]. The involvement of independent researchers attempted to include and integrate varying perspectives in the process of data analysis. Transcripts were analyzed in the original language by Albanian-speaking researchers to enable cultural interpretations that may be diluted during translation. This involved reading and reviewing the transcripts several times while considering the following questions: “Who is telling?” “Where is this happening?” “When did it happen?” “What is happening?” and “Why?” [18]. These questions enabled the analyst to steep the data in meaning [18]. Considering different morphological characteristics of the Albanian and English language, line-by-line analysis was impossible and noncomparable between languages; therefore, sentence-by-sentence initial descriptive open coding was carried out for both Albanian and English transcripts. Once the analysts were fully immersed in the data, coding began by identifying and categorizing all instances of ORIC categories. Operational definitions of ORIC categories were developed to ensure systematic analytical processes, given that more than one researcher was involved in the analysis. The transcripts were read through and all texts that on first impression appeared to represent an ORIC category were highlighted. Highlighted passages were then coded using the predetermined codes. Any

text that could not be categorized with the initial coding scheme was given a new code. Following this phase, the research team reviewed the data for each category to determine whether subcategories were required. Data that could not be coded into an ORIC theory category but that were relevant to organizational readiness was reexamined to describe different types of organizational readiness. Key emerging themes were mapped into a framework that was reviewed and confirmed by all authors [20,23].

Results

Participant Characteristics

Six FGDs were conducted between November and December 2017 in Albania (n=3) and Kosovo (n=3). FGDs included 9-12 participants. In total, 69 professionals participated in this study. Participants represented a spectrum of health professionals working at and in association with CMHCs, with nursing profession accounting for the highest number of participants (25/69, 38%), followed by social workers (13/69, 19%), psychiatrists (13/69, 19%), psychologists (11/69, 16%), general practitioners (GPs) (4/69, 6%), occupational therapist (1/69, 2%), and speech and language therapist (1/69, 2%). GPs external to CMHCs were invited to the FGDs to explore referral routes to CMHCs. Across all sites, 70% (48/69) of the participants were females, and the average age was 41.3 years (range 25-64 years). Managerial positions were held by around 20% of the participants (14/69). On average, years of work experience in mental health was 8 years in Albania and >15 years in Kosovo (see Table 1, for participant characteristics). On average, each FGD took 60 minutes in Albania and 80 minutes in Kosovo. Ahead of the FGDs, participants were asked to rate the following question on a scale of 1-10: “Do you feel the iCBT service delivery will become a normal part of your work?” (a single question that was integrated into their sociodemographic form). On average, both the Albanian and Kosovan sites rated the implementation of iCBT highly, with average scores of 8.42 and 7.48, respectively, out of a score ranging between 1 and 10. Responses for this question were provided by all participants, with the exception of 4 participants from Prishtine and Mitrovice (n=1) and Gjilan (n=3) FGDs.

Table 1. Participant characteristics across Albanian and Kosovan sites.

Characteristics	Albania				Kosovo			
	Tirana	Shkoder	Korce	Total	Prizren	Gjilan	Prishtine and Mitrovica	Total
Participants per site (n)	12	9	12	33	15	10	11	36
Gender^a (n)								
Female	9	7	8	24	9	6	9	24
Male	3	2	4	7	6	4	2	12
Age (years), mean (SD) ^a	34.2 (5.65)	36.1 (8.89)	39.3 (10.03)	35.85 (8.20)	46.47 (8.66)	44.88 (7.85)	46.6 (8.92)	46.12 (8.32)
Years of experience, mean (SD)	7.3 (2.09)	5.9 (3.36)	10.3 (8.49)	8.01 (5.71)	>15 ^b	>15 ^b	>15 ^b	>15 ^b
Managerial position (n)	0	1	1	2	5	3	4	12
Professions (n)								
Psychologists/counsellor	3	2	2	7	1	1	2	4
Psychiatrists	0	2	2	4	3	2	4	9
Nurses	2	4	6	12	6	5	3	14
Social workers/psychosocial counsellor	4	1	2	7	4	1	1	6
General practitioner	1	0	0	1	1	1	1	3
Occupational therapists	1	0	0	1	0	0	0	0
Speech and language therapists	1	0	0	1	N/A	0	0	0

^aThere were missing data for age (n=2) in Gjilan and gender (n=1) in the Prishtine and Mitrovica focus group discussions.

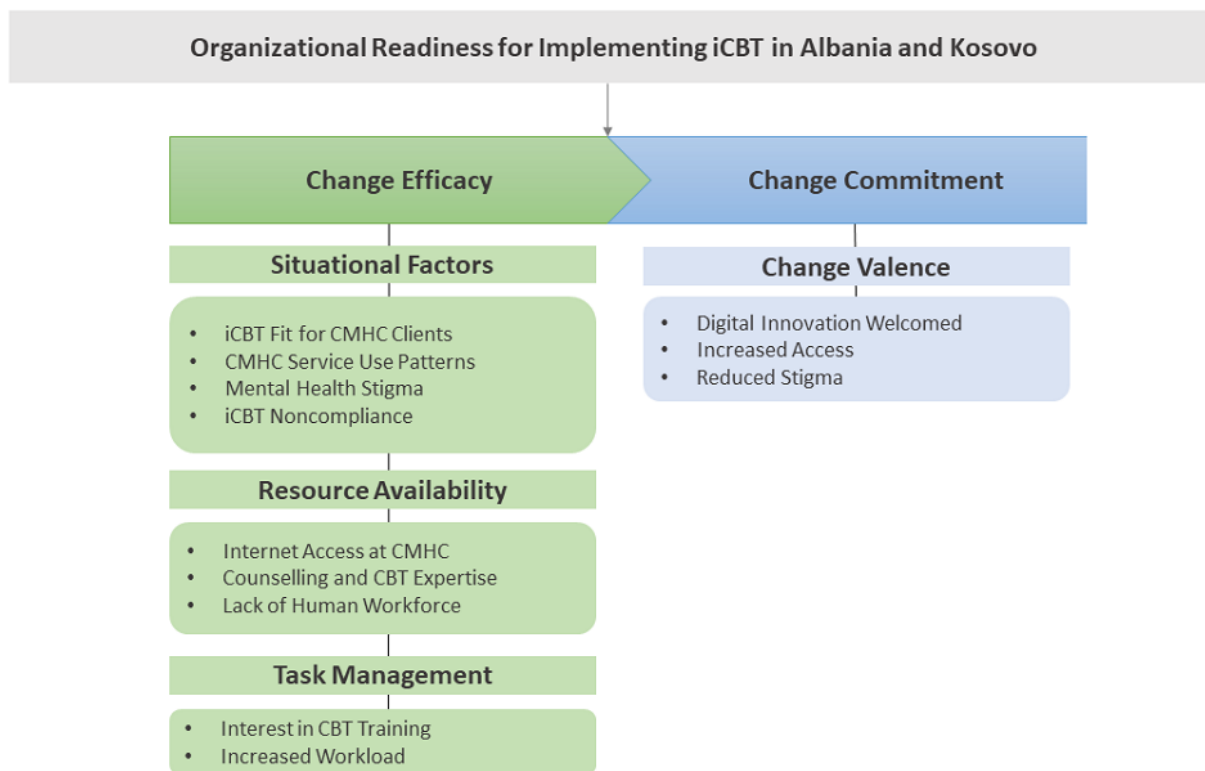
^bYears of experience in Kosovo were captured using 5 brackets: 0-2 years, 3-5 years, 6-10 years, 11-15 years, and >15 years.

Qualitative Framework

A directed content analysis revealed a multifaceted and multilevel conceptual framework of organizational readiness for implementing iCBT across 7 CMHCs in Albania and Kosovo. [Figure 2](#) outlines the key themes that impacted change

efficacy (the belief in the CMHC's collective/individual capabilities to implement iCBT based on existing situational factors, resource management, and task management of services) and change commitment (individual and shared resolve to implement iCBT based on the perceived value to the service) for implementing iCBT in Albania and Kosovo.

Figure 2. Conceptual framework of the factors that influence organizational readiness for implementing internet-based cognitive behavioral therapy in community mental health teams across Albania and Kosovo. CMHC: community mental health center; iCBT: internet-based cognitive behavioral therapy.



Change Efficacy Situational Factors

Situational factors refer to CMHCs members’ judgment and perceptions of contextual factors (eg, beliefs, values, habits) that impact the effective implementation of iCBT. The results of the FGDs revealed that the fit and appropriateness of iCBT to clients accessing CMHCs were perceived as both a barrier and facilitator for implementation, while CMHC client service use patterns, mental health stigma, and concerns around iCBT noncompliance were identified as barriers for implementation.

iCBT Fit for CMHC Clients

FGD participants in Prishtine, Mitrovice, Prizren, Tirana, and Shkoder reported conflicting information about the level of demand for depression-related interventions. One of the participants from Shkoder said:

...Depression is the most common disorder that we see on [a] daily basis. Psychologists provide counselling mostly for depression and other mild mental illness while the rest of the staff, such as the nurses provide treatment for patients with severe or psychotic mental illness that need emergent assistance. Based [on] our statistics, depression has the highest rate in relation to the total number of people who request treatment in our center. [A016]

A mental health professional from the same team elaborated that severe mental illness cases received a disproportionate amount of attention, primarily in response to the pattern of service use in CMHCs that appear to be driven by a desire to access welfare services:

...We don’t treat mild and moderate depression because we are focused on treating patients with severe mental illness, [we] help and support their families and provide social and economic assistance to the patients and their families. [A018]

Considering all clinical groups accessing CMHCs, participants in the Tirana, Prizren, Prishtine, and Mitrovice FGDs reported that there were very few people seeking treatment for mild to moderate depression. They also reported that people who present with depression are typically seen by a GP and thus not treated with psychotherapy. The FGDs in Tirana and Gjilan reported that CMHCs were largely accessed by people with severe mental illness. Not having enough people access services for common mental disorders was perceived as a barrier for the effective utilization of iCBT. Although all GPs reported that they saw large numbers of people with mild to moderate depression, “We have (see) mild and moderate depression which we refer for treatment to a psychiatrist, we don’t prescribe medication without consulting a psychiatrist first” (A007), this was not reflected in the clinical groups accessing CMHCs:

...we treat mostly patients with major mental health disorders, including psychosis such as schizophrenia, mood disorders with psychotic episodes. We have a small number of patients with anxiety and depressive disorders and other mental health. We have few patients with mild and moderate depression and anxiety disorders. [A005]

CMHC Service Use Patterns

Closely related to the perceived inadequate fit of iCBT for the clinical groups accessing the CMHCs, participants in the Tirana

and Shkoder focus groups suggested that there were 2 key patterns for using the CMHC. First, they reported that people generally attended services when their condition was severe and typically discontinued their use of the CMHC when they got better.

...When their (clients') mental health condition gets worse, they come to us for treatment and when they get better, they stop the treatment because they think that they don't need to continue the treatment anymore. [A014]

...We also have patients who come only once for evaluation and a month-long treatment and then they stop coming to our center...They continue to self-medicate for years buying medicines at the pharmacies over the counter and then they show up again. [A003]

Second, participants in the Tirana and Shkoder FGDs also suggested that typical clients were from lower socioeconomic status with severe mental health conditions who accessed the CMHC with the main purpose of applying and qualifying for disability allowance, as CMHCs hold the authority to assess and certify mental health-related disabilities in Albania. Accessing CMHCs therefore appears to function as a means of accessing welfare services:

...That's why we don't have patients with mild or moderate depression only, most of them have chronic or major disorders that come for the application of the reimbursement of the medication they take. [A013]

...The focus of our work is disability assessment and certification.... and we normally follow up those who we certify as disabled, visit them at home and make sure they take their medication... hence these are the issues we are focusing on. [A005]

Participants stated that all CMHCs were situated in major towns, which makes access to services from rural area populations very challenging and that may consequently affect the timely treatment of people with mental illness in the region:

...The problem is related to the access of our services. Those who live in the city have an advantage over those who live in remote areas. They can easily access the service compared to residents who live in rural areas. When patients from remote areas come to us for treatment they have already developed a chronic mental health condition. [A017]

...When they come to us they are in a deteriorated state. Some of them come for treatment after attempting suicide or when they are struggling. [A015]

When prompted further to explore how information on clinical presentations was recorded upon service access, there appeared to be some uncertainty regarding the number of clients and their diagnosis at the CMHC, in which participants suggested that while individual client files were kept, no specific service-wide records were being collected, reported, or analyzed, which is itself a barrier to implementing iCBT.

... We do not have a database. [K2104]

...We don't have numbers of patients being treated based on their diagnosis. [A013]

Mental Health Stigma

FGD participants across both Albania and Kosovo unanimously reported that self and public stigma in relation to mental health and seeking treatment in mental health care services was high. They mentioned that clients did not want to be associated with a label that could lead to social exclusion, which may prevent people from seeking treatment for common mental health conditions.

...I have an example, three of our patients live in the same building and fear that their neighbors might know about their mental health problems. And so, when we do our home visits, we have to wait for one of the patients to close their apartment door after the visit in order to knock at the other patient's door to enter in, in order not to be seen that we have been in their houses despite the fact that they know about each other's mental health issues, and every time we go to their homes they talk about each other. [A004]

Interestingly, as described earlier, the primary service is inundated with individuals with depression as this GP accounts:

...Almost, every day we have similar cases with mild depression but who refuse to go to a psychiatrist. They are afraid and withdrawn because they think that it might be a mental illness problem and get scared of what might happen to them if that would be the case. Meanwhile we give them advise but we would like to have a psychologist to handle this kind of work. [A007]

They do not want to be referred to CMHCs because of the stigma associated with mental health as well as the institutions that provide such support:

...I think that because of the stigma, people suffering from depression avoid seeking help in specialized mental health center such as this one... [A001]

A participant in Shkoder mentioned that it was difficult for participants to accept their diagnosis, leading them to seek different consultations before accepting their condition.

...Usually, they (patients) seek several consultation because it's difficult to accept a mental health disorder. [A007]

Taking this into consideration, it appears as though stigma may be the driving force behind a lack of access to mental health services by people with mild to moderate depression and service user access patterns across Albania and Kosovo.

iCBT Nonadherence

Another perceived barrier expressed by participants in the Tirana FGD concerned adherence to iCBT. Participants in the Tirana FGD expressed that some patients with depression were difficult to engage. While some participants did not see obstacles in engaging adolescents and adults to iCBT, they expressed concerns around engaging older adults with depression.

...Those who suffer from depression will be very difficult to engage through the iCBT intervention. As for the other patients with different diagnosis I think that it won't be a problem. [A005]

...I don't see any obstacles for adolescents, while for adults and elder patients it will be difficult. [A014]

Perceived concerns around confidentiality and difficulty of engaging people with depression appeared to make participants question whether clients would adhere to the iCBT intervention. Participants also expressed concerns around the clients' lack of access to the internet and computers.

...The patients we see are very poor and often they come to us to get the disability benefits because it the only income resource for them. The phone is a more appropriate way of getting in touch with them. [A004]

Participants perceived that clients may have concerns around confidentiality of the iCBT platform, fearing that their information will be accessed by others. Such concerns stem from the pervasive stigma of using mental health services in Albania.

.... I think that confidentiality will be an issue. It will be difficult to persuade our patients that every data that will be put on the platform will be confidential, at the beginning we would need to explain this. [A003]

Resource Availability

Organizational resources broadly refer to the CMHC's existing financial, material, human, and informational resources that are used to make a judgement about implementing and delivering iCBT. The FGDs suggested 2 resource facilitators for implementing iCBT, including appropriate counselling and CBT expertise and internet access, and 1 barrier relating to the lack of available human workforce.

Internet Access at CMHC

Participants across all Albanian FGDs suggested that their CMHCs lacked technological infrastructure and that an investment was required. However, participants in Kosovo reported having a good technological infrastructure in place.

... We have some problems with the internet provider, there are interruptions from time to time. [A021]

Participants suggested that setting up iCBT within the service would be feasible should clients wish to access iCBT through the CMHC.

...Of course. It would be a good idea if the patients receive help from the psychologist while getting access to the iCBT. It's a good thing to have a computer and internet installed in this room and we could help the patients to fill in the worksheets and the patients can come whenever they want in order to make use of the iCBT platform. [A004]

Counselling and CBT Expertise

All FGDs reported having clinical professionals within the team, primarily psychiatrists and psychologists who could be leveraged

to implement iCBT. However, most participants with counselling experience (eg, psychologists, psychiatrists) reported that they did not have the necessary CBT accreditation, with those with knowledge around CBT stating that the training they received was either superficial or did not include a practical component. Moreover, participants suggested that they did not have the general supervision infrastructure to support iCBT.

...Besides to what we have learnt in the university classes we have not been trained in CBT. [A021]

... We had some hours in our psychotherapy module (in University) but it was a short introduction to CBT and we had to read a (book) chapter by ourselves but no practice. [A015]

Participants stated that CBT training could only be accessed privately, requiring clinicians to pay out of their pockets and that training was difficult to access:

... We were notified not long ago about the starting of a training course on CBT, but it was organized in Tirana and the course was very expensive. [A018]

...I try to apply CBT and I face a lot of difficulties... I have had private consultation with the only person who is trained on CBT in the city. [A015]

The team in Prizren also reported using validated (albeit not in Albanian) psychometric scales to screen and monitor the progress of clients, while the FGD in Tirana reported that assessments for mental health were performed by a psychiatrist as a screening but not monitoring measure. This was deemed to be a strength that would be of help in monitoring client outcomes in iCBT.

...[We have used] Hopkins Questionnaire for Measurement of Depression and Anxiety, Beck Depression Scale, Mini International Neuropsychiatric Interview. [K2205]

Lack of Human Workforce

Despite the presence of a small number of clinical staff with limited counselling or CBT experience, participants in most CMHCs reported feeling overstretched owing to expectations to treat clients with heterogeneous clinical presentations and across a wide geographic radius. A lack of adequately trained staff appears to present a barrier to the implementation of iCBT.

...[we see] All disorders, psychotic disorders, personality disorders, depression, anxiety, phobia etc... [K2206]

... We have a geographically heterogeneous patient population (ranging) from the city to the remote villages. [A004]

...1/3 city's center population. We cover 4 or 5 municipal units and its primary health care centers, also some parts of the city's suburb areas. [A013]

...According to the mental health Law, a CMHC should cover 50-150 thousand inhabitants, but in reality we cover 300 thousand inhabitants...threefold to recommended coverage. [A003]

Task Management

Task management refers to the CMHCs' knowledge of task demands by iCBT, while cognitively appraising the match between task demands that are available. The FGD findings suggested that while interest in CBT training was perceived as a facilitator for effectively managing iCBT tasks, an increase in workload was perceived as a barrier.

Interest in CBT Training

FGD participants across Tirana, Prizren, Prishtine, and Mitrovice said that they wished to gain CBT skills. Psychologists, GPs, and psychiatrists expressed that they would like to gain new skills, while social workers and nurses who do not typically deliver CBT expressed an interest in being trained.

...If you involve the social workers and nurses as participants on these trainings, then yes. [A006]

...Such training would be welcomed as we have not had such training and it would be very beneficial. [K2301]

Higher cadre staff such as GPs and psychiatrists were especially keen that staff of a lower cadre such as nurses be trained in CBT, suggesting that it could further enrich the quality of services provided to all clients.

...We work in different ways the counsellor and the nurse give advice to the patient and helps them but none of them work without the permission of [the] doctor. We should train the nurses like most of countries that are free to take the history on its basis... We have 1200 folders and our work is very hard and we want to give service to patients. [K2104]

Increased Workload

Participants in the Shkoder and Tirana FGDs suggested that despite the perceived benefits of introducing iCBT, they anticipated that the project would take a toll on their workload.

...Besides the working benefits that we gain from this project I will have to work extra hours to what I already do here. [A021]

Participants in the Tirana and Shkoder FGDs also suggested that they felt used by previous research projects and enquired if monetary incentives would be available for their additional efforts. They also suggested that clinicians should have the opportunity to decide if they wanted to participate in the trial and implementation in accordance with the additional strains that this may take on their work.

...This will be an extra engagement from my side and I have the right to decide if I want to be involved or not in this project... we understand the aim and structure of the project, we would like to know if there is going to be a monetary incentive/wage for the participants. We have had a lot of people who wanted to implement their projects and we have felt used by them. [A021]

Change Commitment (Change Valence)

Change valence broadly refers to the CMHC member's perceived value of iCBT. FGD data suggested that participants welcomed the innovation of iCBT treatment and valued the role of iCBT for people who are living geographically isolated lives and for reducing stigma associated with accessing mental health services.

Digital Innovation Welcomed

Participants across the Tirana, Prizren, Prishtine, and Mitrovice FGDs appeared to express optimism and high hopes for the use of iCBT, which participants perceived as innovative and potentially leading to improvements in the choice and quality of services.

...We would all be interested about CBT and iCBT training in order to have a different approach with clients. The online intervention would be most welcomed and knowing that this will be the first time that is being implemented in Kosovo we are ready to support it. [K2101]

Increased Access

FGD participants from Prishtine, Mitrovice, Tirana, and Prizren said that online interventions could lead to an increase in service access for people who are residing in geographically isolated regions, living isolated lives, and people who are living in poverty. It was also reported that the implementation of iCBT could remove access barriers related to travel, thereby increasing access to the CMHCs.

...It will be as an open door and relief for depressed patients who need additional psychotherapy. [iCBT] saves client time (travel to psychotherapist) because eg. If fight depression tool can be used in distance. [K2205]

...this online method would be very welcome especially for women who live isolated [lives] for different reasons knowing that technology is used from every person and can be treated online without problems so we as a therapist would be very much pleased and satisfied if we were training in this very important field. [K2101]

Reduced Stigma

While stigma was perceived as a barrier for meaningful exposure to iCBT, FGD participants from Tirana, Prizren, Prishtine, and Mitrovice collectively suggested that the use of iCBT could reduce the stigma associated with mental health care access. For instance, participants reported that clients were concerned about being seen accessing the CMHC, fearing how this might be perceived by their wider community.

...Sometimes, they might come at the treatment center at the same time and pretend to not know each other. [A005]

Participants reported that some of the CMHCs also serve as day centers for long-term community clients, reinforcing that centers mainly support people with severe mental health conditions, thereby increasing the stigma experienced by clients. The

versatility of a web-based CBT intervention could allow clients to engage in treatment for depression at home where they may feel more comfortable and less stigmatized.

...It (iCBT) would be very successful by knowing that stigma is very present, there is always a problem to convince patients to take the right services and this is the reason why this project will have success. For example, if a patient comes to [name of service], they can see patients with serious mental illness here staying in Day Center which increase fear and stigma and if I will show them this web-based tool I'm very sure they will feel more comfortable working from home. [K2206]

Discussion

Principal Results

We conducted a qualitative examination of organizational readiness for implementing an iCBT intervention for people with mild to moderate depression with health care professionals from 7 CMHCs across Albania and Kosovo. Our directed qualitative content analysis revealed 2 overarching themes of organizational readiness for implementing iCBT that aligned with Weiner's [13] ORIC model that we used to interpret the data heuristically. The first was change efficacy, referring to the perception of how possible it would be to implement iCBT in relation to 3 themes, namely, situational factors, resource availability, and task management, which were largely perceived as barriers. The second theme was change commitment, which included 1 subtheme, change valence, in which participants largely expressed that iCBT could result in benefits to their organization and the communities they serve.

Comparison With Prior Work

Several situational factors were identified as barriers for implementing iCBT for depression. Most participants suggested that iCBT was not a good fit for the clinical populations using the CMHCs, which largely consisted of clients with severe mental illness (eg, schizophrenia, mood disorders with psychotic episodes). Moreover, it was reported that other clients accessing the CMHC discontinued their use of the service once their condition was perceived as manageable. These patterns of service use appear to be consistent with the treatment prevalence for psychoses being higher when compared to the treatment prevalence of depression in upper-middle-income countries in contrast to that in high-income countries in which rates are comparable. This demonstrates that people with depression may be underserved in the region [2]. Participants from the Albanian FGD also reported that people with mental health conditions with a low socioeconomic status were more likely to access services as a way of applying for disability allowance and in which access to the CMHC provided a means of accessing social support. While data from the FGDs appear to indicate that stigma is the driving force for the lack of poor uptake by people with low-to-moderate presentation of mental illness, the lack of specialist community mental health services to address common mental disorders may also deter people from accessing services. This is supported by research that perceiving treatment

or services as undesirable can lead to delays in access to treatment [9].

Participants expressed concerns that their clients may fear data privacy breaches and that this may negatively impact access to iCBT. Similarly, a qualitative study exploring the views of people with severe mental illness on digital mental health interventions [27] found that participants expressed concerns about the privacy and confidentiality of their data. Nevertheless, participants also reported that they would still be willing to engage in digital mental health interventions if they felt fully informed and reassured about the use of their data [27]. Increasing transparency around how client data are processed may therefore be helpful in reducing fears around data privacy. While the clinical presentations of people accessing CMHCs appear to be a barrier for implementing and delivering iCBT for depression, some participants (particularly GPs) suggested that they were seeing a high volume of people with symptoms of depression that were hesitant to accept a mental health diagnosis. Participants appeared to link the low uptake of services to the high levels of stigma associated with mental health care in the region. Social stigma of mental illness has been reported globally, not only in Eastern Europe. A review of mental health care systems in Eastern Europe found that experiences of public stigma were higher in countries in Central and Eastern Europe as compared to those in European Union countries [8]. The high social cost of mental illness may therefore negatively impact a person's access to treatment and their mental health trajectory [8]. Remote access to mental health interventions such as iCBT may therefore be critical in reducing the stigma associated with in-person interactions within brick-and-mortar mental health services. A survey on mental health service utilization in Southeastern Europe that included Albania found that perceived stigma led to delays in accessing services as long as 3 months [9]. Taken together, these findings support reports by participants in our study that service access may be negatively impacted by the pervasive stigma associated with mental illness and treatment-seeking behaviors in the region.

Public stigma appears to negatively impact how people access mental health services, requiring significant shifts in public perception around how mental health and treatment access are perceived [28]. Public-facing campaigns such as television programs that raise awareness of mental health treatment or that raise the voices of those experiencing mental health conditions have been put forward as a possible solution for addressing mental health stigma [9]. While evidences regarding the effectiveness of public-facing antistigma interventions in low- and middle-income countries are limited [29], there is some evidence to suggest that they can be effective. A study evaluating the impact of a brief television news report in reducing public stigma associated with schizophrenia in Brazil found that stigma scores were significantly lower in relation to the desire for social distancing and restrictions to the rights of people with schizophrenia after watching the news report vignette and when they were followed up at 3 months [30]. Promising evidence also comes from an antistigma program (Time to Change Global), which aimed to tackle stigma and discrimination toward people with mental health problems in

Ghana and Kenya. The project found significantly positive changes in stigma-related desire for social distancing from people with mental health problems, intent to contact with people with mental health problems, and stigma-related knowledge after a short-term public campaign [31].

Participants reported both facilitators and barriers in relation to resource availability for implementing iCBT. The general technological infrastructure required for iCBT (eg, access to computers, internet at the service) was present in Kosovo but not in Albania where substantial investment was required. Participants said that there were clinical professionals within the team (ie, psychologists and psychiatrists) who either had CBT or counselling skills that would enable them to deliver the intervention. However, when probed further, they explained that their CBT skills were limited and gained by reading book chapters during their university courses, with no clinical practice or supervision. With the exception of 1 psychiatrist in Kosovo who had received formal training, none of the 11 psychologists and 13 psychiatrists had received clinical training on CBT or other psychotherapy approaches, as clinical training and practice are not included in the academic curriculum for Clinical Psychologists in either Albania or Kosovo. Consequently, CBT training was welcomed not only by psychologists and psychiatrists but also by a range of mental health professionals, including those who did not traditionally deliver CBT, such as nurses and social workers. Numerous studies have identified similar barriers in relation to the implementation of iCBT. A qualitative study on the facilitators and barriers of delivering blended CBT (therapist plus computerized program) in the United Kingdom found that the intervention was found to be time consuming, disruptive to usual practice, and overwhelming owing to having low perceived confidence and practice in supporting clients through iCBT [32]. The provision of e-mental health interventions appears to be the simplest part of implementing blended CBT, warranting the development of a toolkit that can enable implementation to be tailored around service needs [33]. Moreover, it is imperative for academic curriculums of Clinical Psychology MSc level courses in Albania and Kosovo to integrate clinical training in their syllabus as well as adopt service-level adjustments that permit staff with the time, resources, and clinical supervision to deliver iCBT in order to increase organizational readiness and facilitate the effective implementation of iCBT in Albania and Kosovo.

Participants also suggested that staff were overstretched by providing services for populations up to threefold more than that recommended by law. Barriers to task management were also associated with a high workload and lack of incentives or compensation to take on additional tasks. The lack of resources and availability of specialist mental health facilities often mean that community mental health services are required to cater for a broader range of people. It has therefore been suggested that transdiagnostic interventions may have wider feasibility and application within low- and middle-income countries [33]. Adopting a versatile iCBT program can be used for a range of different disorders to increase utility in day-to-day practice. Notwithstanding the barriers for implementing iCBT, participants appeared to welcome innovations in the service, perceiving a range of benefits for isolated and stigmatized clients

that enabled greater access to mental health services. A metasynthesis review from high-income countries on the user experience of iCBT found that client participants appreciated the privacy and access that it allowed for people who felt stigmatized by mental health services or those who experienced mobility problems due to a physical disability [34].

Limitations and Strengths

Our study had several limitations. While Albania and Kosovo are located in the same region and have the same language, culture, and values, they are 2 distinct countries with different health care systems that are presented with different challenges and demands. Data for both settings were merged because there were not enough data to conduct a separate analysis for each country. As a result, contextual or nuanced interpretation of the data could not always be generated. We attempted to address this limitation by indicating which sites endorsed different themes. Moreover, the qualitative data analysis did not reveal any conspicuous differences between the sites. Nevertheless, many commonalities were discovered, allowing for greater generalizability to be made. The CMHCs that implemented iCBT mainly provided mental health services for people with severe mental illness because there are little-to-no publicly available psychological services. The mismatch between people who access CMHC services and the target population for iCBT may have negatively impacted perceptions around the feasibility of the implementation of iCBT, even though participants reported that they would value the use of digital innovations in their service. Nonetheless, implementing iCBT within CMHCs was the most feasible option in a region that has limited mental health resources.

This study also had several strengths. Our study is the first to examine organizational readiness for implementing a digital mental health intervention in both Albania and Kosovo. A wide range of professionals participated in the qualitative interviews, representing all professions working at and in association with a CMHC. There was also a good geographical representation of services across both Albania and Kosovo. The findings of our study provide a unique contribution to the literature in relation to the barriers and facilitators for implementing iCBT in these regions and can be used to develop context-specific solutions for implementing iCBT [17].

Three directions for future research are proposed. First, studies should aim to corroborate findings from our study in relation to stigma and perceived patient barriers to accessing the service, with clients using iCBT. Second, a larger sample size should be employed across different regions of Eastern Europe to enable both in-country and cross-country analyses to be conducted. Third, in-depth face-to-face qualitative interviews should also be used to develop a more detailed account of organizational readiness at an individual level [22].

Conclusion

This study aimed to gain a qualitative understanding of organizational readiness for implementing iCBT in CMHCs across Kosovo and Albania. Our study found that almost all participants valued iCBT as a resource for extending the human workforce to meet the needs of stigmatized and geographically

isolated people. However, several mental health system inadequacies such as insufficient resource availability (eg, inadequate technological infrastructure, limited human work force) and task management (eg, lack of CBT expertise and supervision) present challenges for the implementation of iCBT. Moreover, longstanding situational factors pertaining to mental health stigma have shaped the way in which CMHCs are accessed. Interviews suggested that the majority of people accessing services were in need of urgent assistance mentally and financially. However, there was very little representation from people with mild to moderate presentations such as depression, who are as a result rendered as both underserved and unserved, potentially compounding the effect on their mental health presentation and severity. This pattern of service use appears to be driven by the pervasive stigma associated with mental illness to which there is no easy solution and which requires extensive resources and time to address. While iCBT

offers opportunities to decrease stigma associated with engaging patients to mental health interventions, much work is required to increase the treatment-seeking behaviors of people with depression before iCBT can add value to the CMHC. Organizational readiness for implementing iCBT appears to be low, requiring both mental health care systems and situational factors to be addressed before iCBT can optimally be implemented. We propose that such barriers can be addressed through (1) a public-facing campaign that addresses mental health stigma, (2) service-level adjustments that permit staff with the time, resources, and clinical supervision to deliver iCBT in order to increase organizational readiness and facilitate the effective implementation of iCBT in Albania and Kosovo, and (3) establishing a suitable clinical training curriculum for health care professionals to enable them to provide evidence-based treatments for mental health conditions.

Acknowledgments

The authors thank and acknowledge the ImpleMentAll project for supporting this study. The authors also thank the CMHC leads who facilitated recruitment for the study.

Authors' Contributions

All authors contributed to the conception, design, and analyses in this study. ACP conceptualized the study and its design. AD, ACP, CV, and JH contributed to the data analysis plan. ACP developed the topic guides, with contributions from GQ, NF, AM, and SM. Data were collected by ACP, GQ, NF, AM, and SM. Data were translated into English by AM and SM. AD led the data analysis with significant contributions by ACP, CV, JH, AM, NF, and GQ. AD drafted the manuscript with significant contributions by ACP, CV, NF, AM, SM, and GQ.

Conflicts of Interest

None declared.

References

1. Krupchanka D, Winkler P. State of mental healthcare systems in Eastern Europe: do we really understand what is going on? *BJPsych Int* 2016 Nov;13(4):96-99 [FREE Full text] [doi: [10.1192/s205647400001446](https://doi.org/10.1192/s205647400001446)] [Medline: [29093919](https://pubmed.ncbi.nlm.nih.gov/29093919/)]
2. Mental health ATLAS 2017 member state profile: Albania. World Health Organization. URL: https://www.who.int/mental_health/evidence/atlas/profiles-2017/ALB.pdf?ua=1 [accessed 2021-10-20]
3. Long-term sequels of war, social functioning and mental health in Kosovo. Ministry of Health, Kosovo, Kosova Rehabilitation Centre for Torture Victims, World Psychiatric Association. 2006. URL: https://www.ecoi.net/en/file/local/1308565/432_1170782377_final-quantitative-research-report.pdf [accessed 2021-10-01]
4. Fanaj N, Melonashi E. Understanding and Describing PTSD in Kosovo: A Systematic Evidence-Based Review. *SAGE Open* 2017 Oct 12;7(4):215824401773458 [FREE Full text] [doi: [10.1177/2158244017734588](https://doi.org/10.1177/2158244017734588)]
5. Morina N, Rushiti F, Salihu M, Ford J. Psychopathology and well-being in civilian survivors of war seeking treatment: a follow-up study. *Clin Psychol Psychother* 2010;17(2):79-86. [doi: [10.1002/cpp.673](https://doi.org/10.1002/cpp.673)] [Medline: [20041418](https://pubmed.ncbi.nlm.nih.gov/20041418/)]
6. Wang S, Pacolli S, Rushiti F, Rexhaj B, Modvig J. Survivors of war in the Northern Kosovo (II): baseline clinical and functional assessment and lasting effects on the health of a vulnerable population. *Confl Health* 2010 Sep 21;4:16 [FREE Full text] [doi: [10.1186/1752-1505-4-16](https://doi.org/10.1186/1752-1505-4-16)] [Medline: [20858274](https://pubmed.ncbi.nlm.nih.gov/20858274/)]
7. Wang S, Salihu M, Rushiti F, Bala L, Modvig J. Survivors of the war in the Northern Kosovo: violence exposure, risk factors and public health effects of an ethnic conflict. *Confl Health* 2010 May 28;4:11 [FREE Full text] [doi: [10.1186/1752-1505-4-11](https://doi.org/10.1186/1752-1505-4-11)] [Medline: [20509915](https://pubmed.ncbi.nlm.nih.gov/20509915/)]
8. Winkler P, Krupchanka D, Roberts T, Kondratova L, Machû V, Höschl C, et al. A blind spot on the global mental health map: a scoping review of 25 years' development of mental health care for people with severe mental illnesses in central and eastern Europe. *The Lancet Psychiatry* 2017 Aug;4(8):634-642 [FREE Full text] [doi: [10.1016/s2215-0366\(17\)30135-9](https://doi.org/10.1016/s2215-0366(17)30135-9)]
9. Tirintica A, Andjelkovic I, Sota O, Pirlog M, Stoyanova M, Mihai A, et al. Factors that influence access to mental health services in South-Eastern Europe. *Int J Ment Health Syst* 2018;12:75 [FREE Full text] [doi: [10.1186/s13033-018-0255-6](https://doi.org/10.1186/s13033-018-0255-6)] [Medline: [30555528](https://pubmed.ncbi.nlm.nih.gov/30555528/)]

10. Poppleton A, Gire N. mHealth: bridging the mental health gap in central and eastern Europe. *The Lancet Psychiatry* 2017 Oct;4(10):743-744 [FREE Full text] [doi: [10.1016/s2215-0366\(17\)30359-0](https://doi.org/10.1016/s2215-0366(17)30359-0)]
11. Digital 2020: Albania. DataReportal. URL: <https://datareportal.com/reports/digital-2020-albania> [accessed 2020-10-19]
12. Digital 2020: Kosovo. DataReportal. URL: <https://datareportal.com/reports/digital-2020-kosovo> [accessed 2020-09-20]
13. Weiner B. A theory of organizational readiness for change. *Implement Sci* 2009 Oct 19;4:67 [FREE Full text] [doi: [10.1186/1748-5908-4-67](https://doi.org/10.1186/1748-5908-4-67)] [Medline: [19840381](https://pubmed.ncbi.nlm.nih.gov/19840381/)]
14. Feroz A, Kadir MM, Saleem S. Health systems readiness for adopting mhealth interventions for addressing non-communicable diseases in low- and middle-income countries: a current debate. *Glob Health Action* 2018;11(1):1496887 [FREE Full text] [doi: [10.1080/16549716.2018.1496887](https://doi.org/10.1080/16549716.2018.1496887)] [Medline: [30040605](https://pubmed.ncbi.nlm.nih.gov/30040605/)]
15. Mauco KL, Scott RE, Mars M. Validation of an e-health readiness assessment framework for developing countries. *BMC Health Serv Res* 2020 Jun 23;20(1):575 [FREE Full text] [doi: [10.1186/s12913-020-05448-3](https://doi.org/10.1186/s12913-020-05448-3)] [Medline: [32576174](https://pubmed.ncbi.nlm.nih.gov/32576174/)]
16. Wyatt J, Wyatt S. When and how to evaluate health information systems? *Int J Med Inform* 2003 Mar;69(2-3):251-259. [doi: [10.1016/s1386-5056\(02\)00108-9](https://doi.org/10.1016/s1386-5056(02)00108-9)] [Medline: [12810128](https://pubmed.ncbi.nlm.nih.gov/12810128/)]
17. Bührmann L, Schuurmans J, Ruwaard J, Fleuren M, Etzelmüller A, Piera-Jiménez J, ImpleMentAll consortium. Tailored implementation of internet-based cognitive behavioural therapy in the multinational context of the ImpleMentAll project: a study protocol for a stepped wedge cluster randomized trial. *Trials* 2020 Oct 28;21(1):893 [FREE Full text] [doi: [10.1186/s13063-020-04686-4](https://doi.org/10.1186/s13063-020-04686-4)] [Medline: [33115545](https://pubmed.ncbi.nlm.nih.gov/33115545/)]
18. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008 Apr;62(1):107-115. [doi: [10.1111/j.1365-2648.2007.04569.x](https://doi.org/10.1111/j.1365-2648.2007.04569.x)] [Medline: [18352969](https://pubmed.ncbi.nlm.nih.gov/18352969/)]
19. Green J, Thorogood N. *Qualitative Methods for Health Research*. CA, USA: Sage Publications; 2013.
20. Hsieh HF, 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/)]
21. Weiner B, Lewis M, Linnan L. Using organization theory to understand the determinants of effective implementation of worksite health promotion programs. *Health Educ Res* 2009 Apr;24(2):292-305. [doi: [10.1093/her/cyn019](https://doi.org/10.1093/her/cyn019)] [Medline: [18469319](https://pubmed.ncbi.nlm.nih.gov/18469319/)]
22. Shea C, Jacobs S, Esserman D, Bruce K, Weiner B. Organizational readiness for implementing change: a psychometric assessment of a new measure. *Implement Sci* 2014 Jan 10;9:7 [FREE Full text] [doi: [10.1186/1748-5908-9-7](https://doi.org/10.1186/1748-5908-9-7)] [Medline: [24410955](https://pubmed.ncbi.nlm.nih.gov/24410955/)]
23. Assarroudi A, Heshmati Nabavi F, Armat M, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. *J Res Nurs* 2018 Feb;23(1):42-55 [FREE Full text] [doi: [10.1177/1744987117741667](https://doi.org/10.1177/1744987117741667)] [Medline: [34394406](https://pubmed.ncbi.nlm.nih.gov/34394406/)]
24. Thomas E, Magilvy J. Qualitative rigor or research validity in qualitative research. *J Spec Pediatr Nurs* 2011 Apr;16(2):151-155. [doi: [10.1111/j.1744-6155.2011.00283.x](https://doi.org/10.1111/j.1744-6155.2011.00283.x)] [Medline: [21439005](https://pubmed.ncbi.nlm.nih.gov/21439005/)]
25. Graneheim U, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today* 2004 Feb;24(2):105-112 [FREE Full text] [doi: [10.1016/j.nedt.2003.10.001](https://doi.org/10.1016/j.nedt.2003.10.001)]
26. Mays N, Pope C. Rigour and qualitative research. *BMJ* 1995 Jul 08;311(6997):109-112 [FREE Full text] [doi: [10.1136/bmj.311.6997.109](https://doi.org/10.1136/bmj.311.6997.109)] [Medline: [7613363](https://pubmed.ncbi.nlm.nih.gov/7613363/)]
27. Berry N, Lobban F, Bucci S. A qualitative exploration of service user views about using digital health interventions for self-management in severe mental health problems. *BMC Psychiatry* 2019 Jan 21;19(1):35 [FREE Full text] [doi: [10.1186/s12888-018-1979-1](https://doi.org/10.1186/s12888-018-1979-1)] [Medline: [30665384](https://pubmed.ncbi.nlm.nih.gov/30665384/)]
28. Pattyn E, Verhaeghe M, Sercu C, Bracke P. Public stigma and self-stigma: differential association with attitudes toward formal and informal help seeking. *Psychiatr Serv* 2014 Feb 01;65(2):232-238. [doi: [10.1176/appi.ps.201200561](https://doi.org/10.1176/appi.ps.201200561)] [Medline: [24233070](https://pubmed.ncbi.nlm.nih.gov/24233070/)]
29. Clay J, Eaton J, Gronholm P, Semrau M, Votruba N. Core components of mental health stigma reduction interventions in low- and middle-income countries: a systematic review. *Epidemiol Psychiatr Sci* 2020 Sep 04;29:a [FREE Full text] [doi: [10.1017/s2045796020000797](https://doi.org/10.1017/s2045796020000797)]
30. de Oliveira Bosoni N, Busatto Filho G, de Barros DM. The impact of a TV report in schizophrenia stigma reduction: A quasi-experimental study. *Int J Soc Psychiatry* 2017 May;63(3):275-277. [doi: [10.1177/0020764016688713](https://doi.org/10.1177/0020764016688713)] [Medline: [28084154](https://pubmed.ncbi.nlm.nih.gov/28084154/)]
31. Potts L, Henderson C. Evaluation of anti-stigma social marketing campaigns in Ghana and Kenya: Time to Change Global. *BMC Public Health* 2021 May 08;21(1):886 [FREE Full text] [doi: [10.1186/s12889-021-10966-8](https://doi.org/10.1186/s12889-021-10966-8)] [Medline: [33964900](https://pubmed.ncbi.nlm.nih.gov/33964900/)]
32. Cerga-Pashoja A, Doukani A, Gega L, Walke J, Araya R. Added value or added burden? A qualitative investigation of blending internet self-help with face-to-face cognitive behaviour therapy for depression. *Psychother Res* 2020 Nov;30(8):998-1010. [doi: [10.1080/10503307.2020.1720932](https://doi.org/10.1080/10503307.2020.1720932)] [Medline: [32024447](https://pubmed.ncbi.nlm.nih.gov/32024447/)]
33. Bockting CLH, Williams AD, Carswell K, Grech AE. The potential of low-intensity and online interventions for depression in low- and middle-income countries. *Glob Ment Health (Camb)* 2016;3:e25 [FREE Full text] [doi: [10.1017/gmh.2016.21](https://doi.org/10.1017/gmh.2016.21)] [Medline: [28596893](https://pubmed.ncbi.nlm.nih.gov/28596893/)]

34. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. *PLoS One* 2014;9(1):e84323 [FREE Full text] [doi: [10.1371/journal.pone.0084323](https://doi.org/10.1371/journal.pone.0084323)] [Medline: [24465404](https://pubmed.ncbi.nlm.nih.gov/24465404/)]

Abbreviations

CMHC: community mental health center
FGD: focus group discussion
GP: general practitioner
iCBT: internet-based cognitive behavioral therapy
ORIC: organizational readiness for implementing change

Edited by G Eysenbach; submitted 31.03.21; peer-reviewed by L Balcombe; comments to author 06.05.21; revised version received 07.08.21; accepted 19.09.21; published 01.11.21.

Please cite as:

Doukani A, Cerga Pashoja A, Fanaj N, Qirjako G, Meksi A, Mustafa S, Vis C, Hug J

Organizational Readiness for Implementing an Internet-Based Cognitive Behavioral Therapy Intervention for Depression Across Community Mental Health Services in Albania and Kosovo: Directed Qualitative Content Analysis

JMIR Form Res 2021;5(11):e29280

URL: <https://formative.jmir.org/2021/11/e29280>

doi: [10.2196/29280](https://doi.org/10.2196/29280)

PMID: [34723822](https://pubmed.ncbi.nlm.nih.gov/34723822/)

©Asmae Doukani, Arlinda Cerga Pashoja, Naim Fanaj, Gentiana Qirjako, Andia Meksi, Sevim Mustafa, Christiaan Vis, Juliane Hug. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 01.11.2021. 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

Acceptability, Engagement, and Exploratory Outcomes of an Emotional Well-being App: Mixed Methods Preliminary Evaluation and Descriptive Analysis

Amelia Eisenstadt^{1*}, PhD; Shaun Liverpool^{1,2*}, PhD; Athina-Marina Metaxa³, BSc; Roberta Maria Ciuvat⁴, BSc; Courtney Carlsson⁵, MBA

¹Evidence Based Practice Unit, University College London and Anna Freud National Centre for Children and Families, London, United Kingdom

²Faculty of Health, Social Care and Medicine, Edge Hill University, Ormskirk, United Kingdom

³Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

⁴Division of Psychology and Language Sciences, University College London, London, United Kingdom

⁵Paradym, London, United Kingdom

*these authors contributed equally

Corresponding Author:

Amelia Eisenstadt, PhD

Evidence Based Practice Unit

University College London and Anna Freud National Centre for Children and Families

The Kantor Centre of Excellence

4-8 Rodney Street

London, N1 9JH

United Kingdom

Phone: 44 20 7794 2313

Email: mia.eisenstadt@annafreud.org

Abstract

Background: There is growing evidence suggesting that the emotional well-being of the public has been negatively affected in the past year. Consequently, demand for well-being support has increased. Although there is substantial empirical support for mental health apps that target diagnosed conditions, there is less research on emotional well-being apps. Among existing well-being apps, few studies have been conducted on apps that are based on lived experience and those that seek to enhance users' understanding of their emotional patterns. Thus, the acceptability of these novel apps requires further evaluation before upscaling.

Objective: This evaluation aims to describe the acceptability, engagement, and preliminary outcomes of using an app (Paradym) designed to promote emotional well-being and positive mental health.

Methods: This is a pre-post, mixed-methods, single-arm evaluation that is aggregated with digital analytics data. We anonymously collected real-world data on the demographics and well-being of the participants as well as the usability and acceptance of the app using validated questionnaires and open-ended questions. Participants tested the app for a minimum of 2 weeks before completing the follow-up measures. Google Analytics was used to record the level of app engagement. Chi-square and 2-tailed *t* tests were conducted to analyze quantitative data, and a thematic analysis approach was adopted for qualitative data.

Results: A total of 115 participants completed baseline questionnaires, of which 79.1% (91/115) users downloaded the app. The sample was diverse in terms of ethnicity, including 43.4% (50/115) people who self-identified as belonging to minority ethnic groups. Most of the participants were female (78/115, 67.8%) and between the ages of 18 and 25 years (39/115, 33.9%). A total of 34 app users who completed questionnaires at baseline and follow-up provided valuable feedback to inform the future directions of Paradym. Favorable themes emerged describing the app's content, functionality, and underlying principles. Although usability feedback varied across items, a considerable number of participants (22/34, 64%) found that the app was easy to use. Google Analytics revealed that at least 79% (27/34) of people used the app daily. On the basis of preliminary observations, app users experience increased mental well-being. Post hoc analyses indicated that the reduction in depression scores ($t_{33}=-2.16$) and the increase in the well-being measures ($t_{33}=2.87$) were statistically significant. No adverse events were reported during the follow-up period.

Conclusions: The findings of this evaluation are encouraging and document positive preliminary evidence for the Paradym app.

(*JMIR Form Res* 2021;5(11):e31064) doi:[10.2196/31064](https://doi.org/10.2196/31064)

KEYWORDS

smartphone; app; well-being; awareness; mental health; formative; mobile phone

Introduction

Background

Emotional well-being, as an important predictor of human health and longevity, has the potential to reduce the risk of physical and mental health disorders, is well established in the literature [1-4]. Research suggests that emotional well-being has declined in the general population in the past year [5-7]. Therefore, there is an increased need for accessible support services to meet this demand [6]. Evidence suggests that one mechanism for improving levels of emotional well-being is raising one's emotional self-awareness [8]. Furthermore, apps have been identified as a promising mode of delivery based on their accessibility, scalability, and potential to provide anonymous services [9].

Emotional Well-being and Emotional Self-awareness

Emotional well-being is defined as *a component of mental health that includes happiness, interest in life, and satisfaction* [10]. Researchers and policy makers have made the case for the promotion of emotional well-being to advance human health and reduce the costs associated with poor mental well-being and increased risk of mental health disorders [1,3]. Emotional awareness refers to the ability to recognize one's emotions as they are being experienced and communicate them [11]. Low levels of emotional awareness have been associated with poor emotional well-being and an increased risk of mental health disorders, such as depression [8,12,13]. High levels of emotional awareness have been associated with a reduction in depressive symptoms and increased positive affect and emotional regulation [14,15]. There is also evidence suggesting that emotional awareness can enable flexibility of behavioral responses to difficult emotions and contribute to improved relationships with others [15,16].

Emotional Well-being Apps

On the basis of the positive associations between emotional self-awareness and emotional well-being, a growing number of digital interventions have adopted this approach [9,17,18]. These apps offer a route to support users in improving their emotional well-being through their personal digital devices (eg, smartphones and tablets). Therefore, well-being apps can be used at any time, independent of location, with the added benefit that users can use the apps privately [9]. To this end, there has been a growing demand for and development of well-being apps that have accelerated in recent years [19,20]. A systematic review of 52 apps revealed that within the existing range of well-being apps, the apps have aimed to increase users' management of emotions primarily through mindfulness, cognitive behavioral, and mood tracking approaches [21]. For instance, popular meditation and breathing apps, such as

HeadSpace and Calm, have encouraged users to become aware of emotions and manage thoughts and emotions through daily meditation and breathing exercises [22-25].

Other apps have provided strategies from cognitive behavioral therapy, such as challenging negative thoughts, catastrophizing, and facilitating positive reappraisal [26-28]. Furthermore, some apps use mood tracking to enable users to identify emotions or moods and record and monitor these moods as an approach to emotional awareness [28-31]. In response to particular moods, some apps provide suggested activities to manage mood or emotional responses [30]. Although less frequently available, some well-being apps, such as MoodHacker, have been informed by principles of positive psychology that encourage users to identify their strengths and increase their mindful self-awareness [27].

Of the current range of well-being apps, it was noted that apps drew less frequently on lived experiences to explore well-being topics. Similarly, storytelling as a medium to convey psychoeducational concepts has been less studied in the literature [21]. The closest examples to this approach used fictional characters to guide users through the app [32] or real-life testimonials of other people who overcame difficult circumstances [33]. Further, to the best of our knowledge, there are no studied apps that encourage users to go beyond the exploration of their current emotional states, to inquire into their emotional patterns as a means of understanding their personality traits and tendencies over time [21]. Given the rapid development of apps, and relatively fewer studies on apps that use storytelling, lived experience, and emotional patterns, there appears to be a significant gap in this area that could benefit from further exploration. This is also important, given the need to explore wider techniques that could impact levels of engagement with digital interventions [21].

The App Under Evaluation

The Paradym app was developed to support users in increasing their emotional awareness through learning about their emotional patterns to contribute to enhanced emotional well-being, increased self-awareness, and improved life satisfaction. The app uses storytelling as a starting point to introduce users to psychoeducational content that covers key areas of a person's life (ie, love and relationships, body image, work and success, and identity). The app also incorporates concepts drawn from lived experience, which refers to the first-hand personal involvement of the user and the meaning that their past experience brings to current situations [34]. As a starting point, the app was designed for young adults aged 18 years and above. Before further development of the app and to ensure successful implementation, concepts of the Technology Acceptance Model were applied [35]. The model proposes that the assessment of

perceived usefulness and perceived ease of use can determine whether users will engage with the new digital intervention. It is also recommended that user anxiety be monitored, as previous studies identified anxiety as an external variable in the technology acceptance model [36].

The Process of Evaluation

The current evaluation was further guided by the Medical Research Council guidelines, which suggest addressing possible

Textbox 1. Key questions.

Engagement

1. What are the characteristics of users accessing and using the app?
2. What are the users' levels of engagement with the app?
3. Are there differences between completers and noncompleters?

Acceptability

1. What did users like about the app?
2. How usable or useful did users find the app?

Preliminary outcomes

1. Would users' well-being increase after the intervention period?
2. Would users' mental health symptoms (eg, depression) decrease after the intervention period?
3. Would participants experience severe or high levels of anxiety or other negative experiences during the intervention period?

Feasibility

1. Can users' feedback be used to inform upgrades to the app?
2. Would it be feasible to conduct further evaluations of the app using current recruitment strategies and outcome measures?

uncertainties identified during the development of new interventions [37]. Therefore, in this evaluation, we aimed to describe the acceptability, engagement, and preliminary outcomes of using an app (Paradym) designed to promote emotional well-being by adopting novel approaches. At this formative stage, the paper also aimed to record and report any negative consequences of using the app. To address these aims, the following key questions were addressed (Textbox 1).

Methods

Design

The current evaluation adopted a mixed-methods study design with a pre-post single-arm approach aggregate with digital analytics data. We adhered to the Consolidated Standards of Reporting Trials guidelines [38] where applicable, and registered the user testing protocol on the Open Science Framework a priori [39]. The only change in protocol was the decision to not collect data on anxiety at baseline. As the purpose of the app was not specifically designed to address mental health symptoms (eg, anxiety), to minimize the burden of completing lengthy questionnaires, we only collected data from the anxiety measure at follow-up. However, this was partly in line with our research question to explore whether the use of the app introduces high levels of anxiety.

Participants

Real-world (ie, in-the-wild) data were collected from an international pool of potential users. User testing was advertised on Facebook campaigns and through social media over a 6-week period between May and June 2021. Anyone coming in contact with information about the user testing had the opportunity to participate in user testing. To be eligible for the user testing, participants needed to be over the age of 18 years and not have a diagnosable mental health condition.

Procedure

Participants were enrolled in the user testing through a link that was provided on the web advertisement and on the Paradym website. The link provided access to web-based questionnaires for the baseline assessment and instructions on how to download the app from the Apple app store for iPhone users or Google Play for Android users. An automated email was sent to participants 2 weeks after completing the first battery of questionnaires for participants to complete the follow-up questionnaires. The questionnaires at time 2 included measures completed at time 1 and additional measures to capture the acceptability of the app and the users' experience. Participants who completed questionnaires at both time points were entered into a free prize draw with the opportunity to win 1 of 3 US \$25, US \$50, and US \$100 Amazon Gift Cards.

Intervention Description

Overview

Paradym is an app aimed at supporting emotional well-being within the general population and targets adults over 18 years of age. It aims to provide users with support to develop greater emotional awareness and enhanced life satisfaction. Paradym was initially a crowdfunded project in the early stages of its development specifically to ensure that commercial interests were not taken into account in the early development of the intervention. The app provides a low-cost self-guided program

with the aim of supporting users' levels of emotional awareness and emotional well-being. Paradym is a standalone app, but it can also be used in combination with web-based group coaching sessions. However, for this evaluation, participants only accessed the standalone app.

Paradym was designed drawing on evidence-based strategies and designed in response to key topics that users identified as having such support as success, body, identity, love, and relationships. Clinical psychologists, coaches, and researchers were involved in the development of the content and the selection and review of psychoeducational content, suitable evidence-based strategies, and exercises (eg, journaling and recording personal notes).

Although an integrative theoretical approach underpins Paradym, key therapeutic theories such as acceptance and commitment therapy [40] and schema therapy [41] informed the lived experience and storytelling activities. Following consultation with users, further theory-informed strategies were applied to produce content that users indicated were important. Paradym comprises the following key strategies:

Digital Lessons

Psychoeducational digital lessons are structured into 5 pillars: aware, success, love, identity, and body. These topics were chosen based on the results of an end-user consultation conducted early in the development phase. Examples of psychoeducational content include an introduction to developing emotional awareness [42], supporting users with the identification and understanding of their emotions [43], and identifying and strengthening values [40].

Each digital lesson begins with an explanation of the psychoeducational concept based on lived experiences through personal storytelling. Storytelling is expected to help make the psychological content relatable, foster perspective taking, and support identification with the storyteller [44-46]. Storytelling has been found to boost engagement in both psychological interventions and apps [47]. Digital lessons are provided as chapters for reading, audio for listening, and video to watch speakers present the content to support different learning styles.

Emotional Patterns

At the end of each digital lesson, the user is asked to identify their own emotional pattern in relation to the app's content, which draws on some concepts from schema therapy [41]. Its purpose is to enable the user to gain greater emotional awareness. To achieve this, the app introduces users to a range of emotional patterns and asks users to reflect on their own emotional patterns, and which pattern they perceive themselves to correspond to at the time of completing the lesson. Participants can then revisit specific emotional patterns and change them over time.

Reflections

Reflections are provided to the users via push notifications to prompt engagement with the app through a new daily reflection. Research shows that there are many benefits of reflection, such as supporting determination with tasks despite a stressful context

[48], and that reflection may reduce rumination, which is linked with internalizing difficulties and interpersonal conflict [49].

Furthermore, notifications have been found to increase app engagement. Several studies have indicated that the use of mobile phone app interventions delivering psychological content must be combined with the active engagement of users [50,51].

Data Collection Tools

Demographics

To understand the demographic profile of users accessing and using the interventions, participants were asked to provide anonymous information about their age, gender, ethnicity, and country of residence. Age was captured using age groupings of approximately 10 years (eg, 18-25 years, 26-35 years, 36-45 years). To allow inclusivity, participants were asked to enter their gender and ethnicity using free text. Participants also selected their geographic location (eg, city, state, country) from a global list. For analysis, ethnicity was dichotomized as White or minority ethnic groups. Age was recategorized to represent 45+ years, as there were fewer persons in each age category above 45 years. Location was then categorized as the United Kingdom, United States, and others to represent most responses.

Emotional Well-being and Mental Health Factors

Overview

The outcome measures used to assess well-being were the World Health Organization-Five Well-Being Index (WHO-5) [52], the Satisfaction With Life Scale (SWLS) [53], the Emotional Self-Awareness Scale-Revised (ESAS-R) [8], the Patient Health Questionnaire-9 (PHQ-9) [54], and the Generalized Anxiety Disorder-7 (GAD-7) [55]. WHO-5, SWLS, ESAS-R, and PHQ-9 were administered pre and post app use. The GAD-7 was administered post app use alongside the System Usability Scale (SUS) [56] and the semistructured acceptability questionnaire.

Well-being Measure

The WHO-5 allows a brief assessment (under 1 minute) of well-being over a 2-week period [52]. Individuals are asked to indicate for each of the 5 statements how they felt over the past 2 weeks using a 6-point Likert scale ranging from 0=*at no time* to 5=*all of the time*. The WHO-5 is derived from a 28-item version based on items from the Zung scales for depression, distress, and anxiety, as well as from the General Health Questionnaire and the Psychological General Well-Being Scale [57]. The WHO-5 has been validated as a measure of depression in both adolescents and older adults, with high measurement invariance [58]. A high score indicated a high level of well-being.

Satisfaction With Life Measure

The SWLS [53] is a short 5-item scale designed to measure the global cognitive assessment of satisfaction with one's life. The estimated time for completion of SWLS has been reported to be approximately 1 minute. The SWLS has been shown to have very high construct validity, with Cronbach $\alpha=.85-.87$ [59] and moderately high reliability (Cronbach $\alpha=.78$) [60]. A high score obtained from the SWLS indicates a high level of life satisfaction.

Emotional Self-awareness Measure

The ESAS-R [8] is a 30-item scale, and all items are rated on a 5-point Likert scale ranging from 0 to 5 (0=never, 1=very little, 2=sometimes, 3=often, 4=a lot). The subscales ranged from 0 to 20. The total scale ranged from 0 to 132. Subscales included recognition, identification, communication, and contextualization. The ESAS-R has been shown to have high validity and reliability (Cronbach $\alpha=0.83-0.90$) [30]. A high score on this measure indicates a high level of emotional self-awareness.

Depression Measure

The Patient Health Questionnaire (PHQ-9) is a depression scale that scores each of the 9 *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* criteria as 0 (not at all) to 3 (nearly every day). The PHQ-9 has been validated for use in primary care [45]. It is not a screening tool for depression, but can monitor the severity of symptoms and response to treatment. Scores over 10 have good sensitivity (88%) and specificity (88%) for the diagnosis of major depression by interview. It has high internal reliability, and Cronbach $\alpha=.89$ [61]. The construct validity of the PHQ-9 is also high for community and clinical samples [62]. A high score obtained from the PHQ-9 is indicative of a severe level of depression.

Anxiety Measure

The GAD-7 is a 7-item, brief clinical measure that assesses the presence and severity of generalized anxiety disorder. The self-report scale asks how often during the last 2 weeks individuals experienced symptoms of generalized anxiety disorder. Total scores range from 0 to 21, with cutoff scores of 5, 10, and 15 indicating mild, moderate, and severe anxiety, respectively. Increasing scores on the GAD-7 are strongly associated with greater functional impairment in real-world settings. Sample items are rated from 0 (not at all) to 3 (nearly every day). Scores over 10 have good sensitivity (89%) and specificity (82%) for diagnosis of generalized anxiety disorder by interview, and the scale has high internal reliability, as suggested by a Cronbach $\alpha=.92$ [55]. The scale has been widely used and considered a valid and reliable screening tool in previous research, presenting good reliability, factorial, and concurrent validity [63]. A high GAD-7 score is an indicator of severe anxiety.

Acceptability and Usability

System Usability Scale

The SUS is a 10-item scale with all items measured on a 5-point Likert scale ranging from 1 to 5 (1=strongly disagree and 5=strongly agree) with total scores ranging from 0 to 100. This widely used scale in technology provides an overview of the subjective assessments of usability [56]. The SUS has been shown to have excellent construct validity and reliability (0.81-0.94) [64,65]. A high score on the SUS indicates a high level of usability, with a score above 68 considered above average.

Semistructured Questions

In addition to the SUS, we aimed to capture information about the acceptability of the app by identifying facilitators and

barriers to using the app. For the open-ended questions, the following questions were adapted from a study by Kern et al [66] to fit the current evaluation:

1. Why would you use (Paradym)?
2. Why would you not use (Paradym)?
3. Explain why (or why not) you would prefer to use (Paradym) to seeing a mental health professional?

Adverse Events

To capture any side effects of using the app, participants were provided with the email address for the research team and requested to contact a member of the research team if they experienced any distress or discomfort or experienced any other issue because of their use. Adverse events (ie, side effects) were identified and recorded as any untoward medical or clinical occurrence, which did not necessarily have a causal relationship with the intervention. Any adverse events arising during the evaluation period were assessed for severity, causality, seriousness, and expectedness (ie, relating to the information provided by the app).

Data Analysis

Descriptive statistics were calculated for participant characteristics at baseline (preintervention), and Google Analytics estimates were used to report app engagement. The main focus was on descriptive data to address the aims of the evaluation. However, some exploratory significance tests were conducted on within-group mean differences at the 2 time points (ie, baseline and follow-up) on mental health and well-being measures. Paired samples, 2-tailed *t* tests were conducted on continuous data, and Chi-square tests were used for categorical variables in SPSS [67]. Qualitative data from the semistructured open-ended questions were captured in a questionnaire format, coded in categories, and analyzed using the steps outlined for thematic analysis [68]. The data were coded line by line and then clustered into provisional themes. Each candidate theme was reviewed by going back to the coded data to check whether the meaning was accurate. Next, the uncoded datasets were revised to ensure that the data were relevant to the research question. Finally, the provisional theme titles and the codes that they contained were checked to confirm that they remained relevant and accurate. Any discrepancies that arose were discussed at team meetings and only included after a consensus was reached.

Ethical Considerations

We were guided by the research ethical principles for ensuring the rights, safety, dignity, and well-being of the participants in this evaluation were protected [69]. All data were collected remotely from a nonclinical population, and participants voluntarily provided informed consent for their anonymized deidentifiable data to be used for research purposes. The research team's email address was available for participants to contact us if they experienced any discomfort because of using the app. The researchers were prepared to signpost users to relevant well-being support and resources, if required. Owing to the formative nature of the evaluation, findings were not expected to be generalizable or transferable in any way. We also consulted other researchers affiliated with academic

institutions and practitioners known to the study team who reviewed the evaluation protocol and provided oversight with respect to research ethical guidance. In light of the above, this evaluation was viewed as user testing or service evaluation and therefore was exempt from formal ethical approval by the Health Research Authority ([Multimedia Appendix 1](#)).

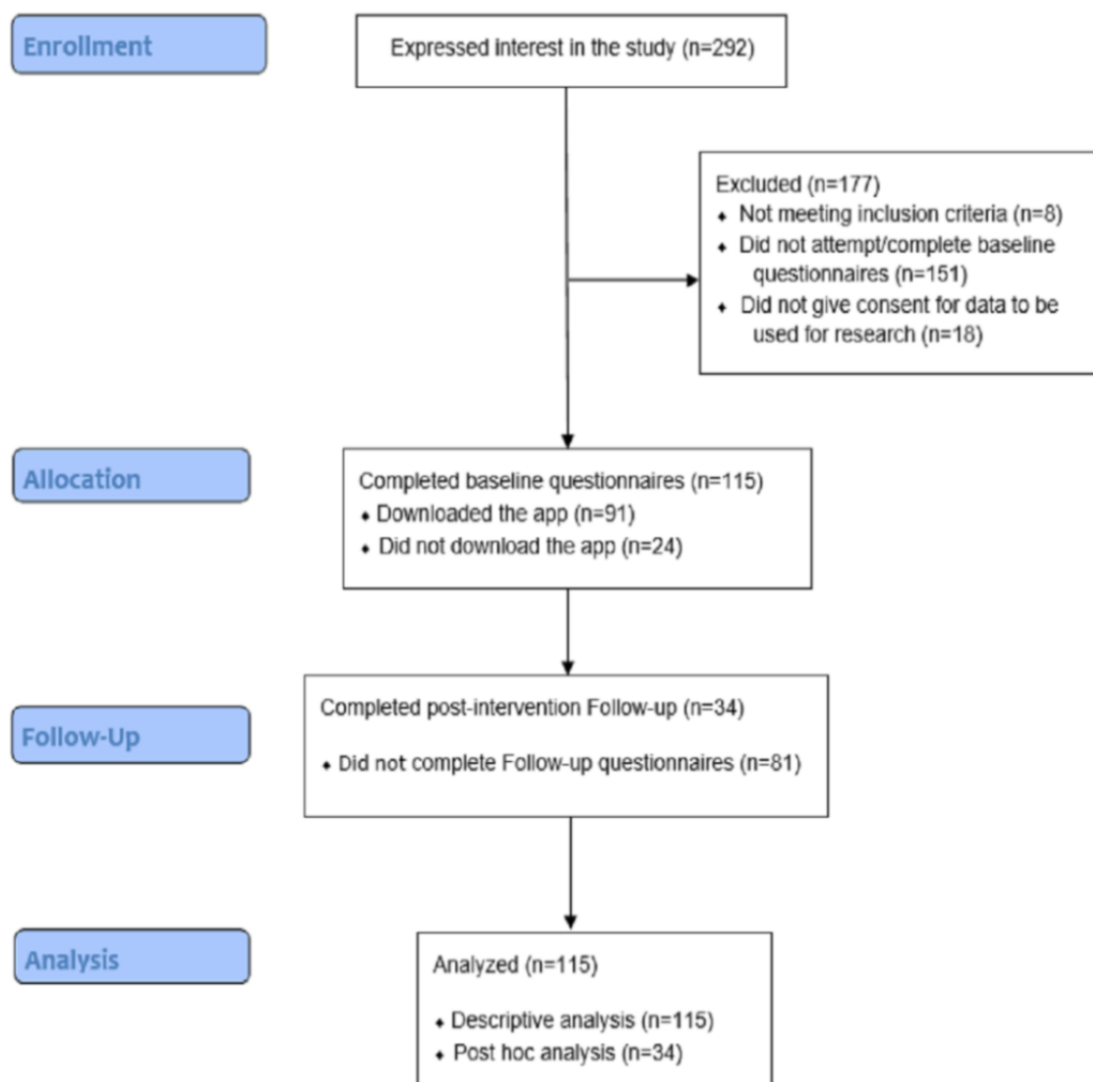
No formal ethical approval was required for similar acceptability studies [30,70,71]. As explained by Ahtinen et al [72], ethics committee approval was not acquired, as the study was deemed to involve minimal risk and the focus was on studying mainly user experiences. In addition, the present evaluation does not involve a clinical population, like that of Ly et al [73]: “Since this pilot trial involved a non-clinical population, it was considered exempt from registration in a public trials registry.”

Results

Overview

Questions about engagement, acceptability, and preliminary outcomes (1-8) are addressed in the results section, and

Figure 1. Flowchart of participants' progress through the evaluation period.



questions on feasibility (9,10) are addressed in the *Future Directions* subsection of the *Discussion* section.

Descriptive Data

During the 2 weeks, the user testing was advertised, and 292 participants expressed interest as part of the preliminary evaluation. One hundred fifty-one individuals either did not attempt or did not complete the baseline questionnaires and therefore did not reach the relevant page to be able to download the app. However, data from 18 participants were not considered, as they did not provide consent to process their information for research purposes. Data from 8 participants were excluded as they indicated that they were not eligible for the evaluation based on mental health diagnoses. Consequently, data from 115 participants (ie, completing baseline questionnaires and providing consent) were included in the evaluation. Of these, 91 participants downloaded the app, and 34 of the participants completed questionnaires at both baseline and follow-up and were included in the final analysis. [Figure 1](#) depicts the flow of participants throughout the evaluation period.

Participant Characteristics

At baseline (n=115), the participants were a self-identified diverse sample with 43.4% (50/115) belonging to minority ethnic groups, with most being female (78/115, 67.8%). Most of the samples (80/115, 69.5%) were from different cities across the United States. Participants' ages ranged from 18 to 65+ years, with the mode being 18 to 25 years. Table 1 provides

further details of the sample. As for mental health and well-being characteristics, the participants were generally *dissatisfied* with their life (mean 20.16, SD 6.24) and had *moderate* symptoms of low mood or depression (PHQ-9; mean 10.35, SD 6.65). Participants also reported *average* scores on emotional self-awareness (ESAS; mean 56.95, SD 13.94) and general well-being (WHO-5; mean 17.22, SD 4.96).

Table 1. Demographic information for the participants at baseline and follow-up.

Characteristics	Participant, n (%)	
	Baseline (n=115)	Follow-up (n=34)
Gender		
Male	29 (25.2)	9 (26.5)
Female	79 (68.6)	23 (67.6)
Transgender	2 (1.7)	0 (0)
Other or prefer not to answer	7 (6)	2 (5.9)
Age (years)		
18-25	39 (33.9)	13 (38.2)
26-30	21 (18.2)	3 (8.8)
31-35	15 (13)	0 (0)
36-40	15 (13)	5 (14.7)
41-45	7 (6)	4 (11.8)
46-50	4 (3.4)	0 (0)
51+	15 (13)	9 (26.5)
Ethnicity or race		
White	66 (57.3)	24 (70.6)
Black	12 (10.4)	2 (5.9)
Hispanic	4 (3.4)	0 (0)
Asian	20 (17.3)	6 (17.6)
Indian	1 (0.8)	0 (0)
Asian Americans and Pacific Islanders	1 (0.8)	1 (2.9)
Jewish	2 (1.7)	0 (0)
Mixed	10 (8.6)	1 (2.9)
Country		
United States	80 (69.5)	23 (67.6)
United Kingdom	32 (27.8)	10 (29.4)
Other	1 (0.8)	1 (2.9)
Did not answer	3 (2.6)	0 (0)

Engagement With the Intervention

App use data were made anonymous to comply with the General Data Protection Regulation and research ethical guidelines. Ninety-one users downloaded the app; therefore, it was not possible to separate the data of users who used the app only and those that used the app and completed the final battery of measures. During the evaluation period, all users logged into the app at least once. The average use across all participants during the evaluation period was 23.53 minutes. The reflections

received 215 views. Users returned to the app an average of 4.5 times with each session lasting an average of 5.22 minutes. Overall, the Daily Active Use Monthly Active Use ratio (ie, the proportion of users that engage with the app in a single day, calculated by dividing the number of daily users by the number of monthly users) was 80% according to Google Analytics.

Attrition or Dropout

In terms of attrition, 70.4% (81/115) users did not complete follow-up questionnaires. The remaining 29.5% (34/115)

participants completed the questionnaires at the 2 required time points. Overall, there were no significant differences in demographic or mental health or well-being characteristics. However, based on descriptive data, participants who completed the evaluation had slightly lower mean scores at baseline on all outcome measures including SWLS (18.94 vs 20.67), WHO-5 (16.18 vs 17.65), ESAS-R (55.95 vs 57.37) and PHQ-9 (10.12 vs 10.45).

Acceptability

Thematic analysis of the qualitative data led to the development of 3 key themes. Participants provided favorable feedback describing acceptance of *the app's content*, *the app's functionality*, and *the app's underpinning principle*. Owing to the interaction of these themes, users' feedback indicated that there is an element of app versatility, with users describing the use of the app at different times of the day. Some described it as part of their morning routine or used the Paradym app in conjunction with other well-being and health apps.

The App's Content

In this first theme, users commented on the content of the app, expressing positive evaluations of the app's chapters, reflections, and exercises, and reported that listening to or reading the content helped them with their understanding of themselves and their mental health. For example, one user expressed, "[Because of the app, I am] studying myself more closely which has benefited my mental health" (Female, 51+ years).

Other users recognized the value of Paradym as a form of additional support separate from that of a professional. They said that Paradym could assist with self-help, whereas their mental health professionals could help with the application of the in-app content. For example, one user expressed, "Paradym allows me to individually learn and apply it onto myself while training myself also. Seeing a mental health professional could allow them to help me directly" (Female, 18-25 years).

The App's Functionality

In this second theme, users reported that the combination of videos and text as aspects of the app's functionality was helpful when working through the exercises as a way of building knowledge. For example, one user said, "The videos were actually pretty engaging when I got around to watching them, and I did feel like I gleaned some insights and self-awareness from that" (Gender not reported, 18-25 years).

Others reported valuing the app as a chance to reflect and understand emotions and found the daily reflection notifications useful: "[A chance to] reflect on my emotions and life. Works best with the daily prompts" (Female, 18-25 years).

The App's Underpinning Principle

In the third theme, users commented on the app's underlying principle. Participants reported benefits from building increased awareness of their emotional and behavioral patterns, moods, and having an opportunity to explore and record them. For example, one participant said, "[The app allowed...] building consistency and having a place to collect my thoughts about my mood and behavioral patterns" (Female, 26-30 years).

Similarly, another user reported that developing self-awareness was important and linked to their personal aspirations. For example, one participant reported, "[I would use Paradym...] to discover bits about myself and become the person I knew I was capable of being" (Male, 41-45 years).

Usability Analysis

The mean usability score was almost 60 (mean 59.77, SD 23.65) out of a possible 100. Participant responses varied across the items on the SUS. On the learnability subscale, over 91% (31/34) of the sample were neutral or disagreed that they would need help to use the app (Item 4), and over 82% (28/34) were neutral or agreed that they could learn to use the app quickly (Item 7). In the same vein, 74% (25/34) of users reported feeling confident enough to use the app (Item 9). [Table 2](#) provides further details on the participants' usability experience.

Table 2. Systems usability scale self-report of items post intervention.

Number	Item	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Missing data, n (%)
1	I think that I would like to use Paradym frequently	14 (41)	9 (26)	11 (33)	0 (0)
2	I found Paradym unnecessarily complex	12 (35)	6 (18)	16 (47)	0 (0)
3	I thought Paradym was easy to use	14 (41)	11 (33)	9 (26)	0 (0)
4	I think that I would need the support of a technical person to be able to use Paradym	3 (9)	6 (18)	25 (73)	0 (0)
5	I found the various functions in this app were well integrated	15 (42)	12 (35)	7 (23)	0 (0)
6	I thought there was too much inconsistency in Paradym	7 (20.6)	9 (26.5)	17 (50)	1 (2.94)
7	I would imagine that most people would learn to use Paradym very quickly	22 (64.7)	6 (17.6)	5 (14.7)	1 (2.94)
8	I found Paradym very cumbersome to use	12 (35.3)	7 (20.6)	14 (41.2)	1 (2.94)
9	I felt very confident using Paradym	17 (50)	8 (23.5)	9 (26.5)	0 (0)
10	I needed to learn a lot of things before I could get going with Paradym	7 (20.6)	8 (23.5)	17 (50)	2 (5.88)

Adverse Events

Regarding possible side effects or adverse events, no reports from the participants were received during the evaluation period. Therefore, the research team was unaware of any harmful outcomes that could result from using the app.

Preliminary Efficacy

On the basis of the available data from the participants who completed the evaluation, a descriptive analysis showed that

mean scores slightly improved on the SWLS (+1.56), the WHO-5 (+2.03), and the ESAS (+1.75). However, the increase was statistically significant only for the WHO-5 score ($t_{33}=2.87$). In addition, participants' depression scores also decreased over time (-1.56) and were found to be statistically significant ($t_{33}=-2.16$). Participants' mean anxiety levels were within the mild category of about 8.24 (SD 6.59) during the intervention period. Table 3 provides further descriptions of the sample ($n=34$) included in the final analyses.

Table 3. Descriptive statistics for each outcome measure at baseline and follow-up.

Outcome measure	Baseline, mean (SD)	Follow-up, mean (SD)	Coefficient (95% CI)	Cohen <i>d</i>	<i>P</i> value
SWLS ^a	18.94 (5.76)	20.50 (7.03)	0.73 (-0.11 to 0.85)	0.37	.07
WHO-5 ^b	16.18 (4.70)	18.21 (4.94)	0.64 (0.02 to 0.99)	0.51	.007
ESAS-R ^c	55.95 (12.50)	57.70 (14.27)	0.50 (-0.34 to 0.62)	0.14	.57
PHQ-9 ^d	10.12 (6.25)	8.56 (6.35)	0.78 (-0.85 to 0.11)	-0.38	.04
GAD-7 ^e	— ^f	8.24 (6.59)	—	—	—

^aSWLS: Satisfaction with Life Scale.

^bWHO-5: World Health Organization-Five Well-Being Index.

^cESAS-R: Emotional Self-Awareness Scale-Revised.

^dPHQ-9: Patient Health Questionnaire-9.

^eGAD-7: Generalized Anxiety Disorder-7.

^fNot available.

Discussion

Principal Findings

The aim of this user testing and initial service evaluation was to describe the acceptability, engagement, and preliminary outcomes of using a new version of a well-being app. The findings indicated that users who downloaded the app ($n=91$) generally accessed and used the app during the intervention period for an average of 5.22 minutes with an average of 4.5 sessions per user over the 2-week period. A fair number of participants who completed the post assessment (10/34, 29.4%) engaged with the app content in all 5 pillars. Despite the observed level of engagement (mean 23.53 minutes), there was high attrition (70.4%) during the evaluation period. Although there were no significant differences between the users who completed the evaluation and those who dropped out, the mean scores for mental health and well-being measures were higher in those who dropped out. It is possible that people who experienced more psychosocial difficulties found it difficult to engage with the app or to respond to the outcome measures [74]. Mental health apps generally have high dropout rates, and some studies have found that once users learn a skill or knowledge from a particular app, they stop using it [75]. It is also conceivable that, for users with high life satisfaction, they may have had less need for the knowledge provided in the app or felt they had obtained what they needed before the end of the 2-week period.

Participants who completed the evaluation provided favorable feedback in terms of the app's content, functionality, and underlying principles. Although usability feedback varied across

items of the SUS, a high percentage of participants found that it was easy to use the app. This can be attributed to evidence suggesting the importance of specific features that could influence users' experiences. For example, younger digital users are more likely to engage with interventions that have features such as videos, limited text, ability to personalize, ability to connect with others, and options to receive text message reminders [76]. Moreover, the promotion of increased self-awareness is becoming popular among young people in their early twenties [77]. Therefore, as emotional awareness is one of the main purposes of Paradym, and a fair amount of content is presented via videos or audio, users could have responded particularly favorably. Similarly, Paradym includes customization features, exercises, and text materials. Contrary to Liverpool et al [76], users in this evaluation responded favorably to text-based content (as per the digital lessons). Further research is needed to determine whether specific features may have more or less impact on use among various age groups.

In addition, based on descriptive data, all mental health and well-being scores improved. The change in well-being (WHO-5) and depression (PHQ-9) were statistically significant, indicating that the use of the app could potentially improve some symptoms related to poor well-being. However, as this was not a controlled study, it is not possible to make any causal claims or explicitly attribute the findings to the use of the app. Nonetheless, this is the first study to evaluate Paradym; therefore, these preliminary findings can be viewed as positive and warrant further investigation. The other outcome measures did not yield statistically significant results. This could be because of the small sample size ($n=34$) or other methodological issues, such as the 2-week intervention period. It could be that factors

associated with life satisfaction and emotional awareness require longer use periods. It was also noted that anxiety levels of the participants were mild during the intervention period, which at a minimum could indicate that the Paradym app did not induce any unnecessarily high levels of anxiety.

Comparison to Other Findings

In terms of the demographic profile of the sample, the evaluation captured data from a diverse sample (50/115, 43.5%, belong to minority ethnic groups). This is generally uncommon in mental health app studies [22] and therefore provides valuable insights from underserved and underrepresented populations. Although we hoped to collect data from a larger sample, the total sample size and high attrition observed in this evaluation appear to be common when evaluating digital interventions [78]. Despite these challenges, the findings are in line with the existing literature, indicating the potential usefulness of apps to support positive mental health and well-being. Our findings suggest a significant increase in well-being, which has also been observed in similar studies [22,24]. A decrease in depression symptoms has also been reported in other studies that evaluated mental health app use [31,79-81]. Similar themes and statistics have also been reported for user engagement and acceptance [82,83]. However, in terms of study design, this evaluation obtained both quantitative and qualitative data, including objective engagement data, to fully capture users' feedback about the app. This evaluation also demonstrated that it was possible to conduct an app evaluation within a 6-week period, whereas previous studies reported slightly longer evaluation periods of 8 weeks [84] or more [85]. When compared with recent systematic reviews and meta analyses, the effect sizes in our evaluation fell within the previously observed range (eg, 0.10-0.57) for depression ($k=12$ [86]; $k=18$ [87]). Furthermore, for well-being, the effect size in this evaluation (0.51) was above the previously observed range ($k=5$; effect size range: 0.14-0.45) [88].

In terms of the app itself, based on available descriptions for similar apps [89], Paradym, to the best of our knowledge, may be one of the first apps that focuses on improving emotional awareness and emotional well-being for the general population through the medium of identifying emotional patterns, and integrates many psychological theories. The module-based content is similar to that of other apps [90-92], but the focus on users identifying and selecting a range of emotional patterns to build an overall profile appears to be unique to Paradym. This is an important finding, as preliminary feedback suggests that some users responded positively to this feature.

Strengths and Limitations

The main strength of this evaluation is its ability to collect and analyze data from a diverse sample of app users. At this preliminary stage of the app's evaluation, the findings already suggest some evidence of promise. These include significant impacts on mental health and well-being and no reports of adverse events, in addition to positive feedback on use and engagement. Another strength is its ability to recruit an acceptable sample size and carry out the evaluation within a 6-week period. This could indicate a need on behalf of the users and demonstrate to other app developers and researchers that it

is feasible to conduct ongoing evaluations during rapid prototyping phases.

Despite its notable strengths, this evaluation also has some limitations. Obvious limitations include the single-arm pre-post evaluation design, which limits the ability to perform comparative analyses. The moderate sample size ($n=34$) has limited potential to carry out subgroup analyses (eg, males). Studies with larger sample sizes and subgroup analyses could provide valuable information for future versions of the app. The high attrition rate (70.4%), which is in line with results reported by other studies, is a point of concern, and future evaluations of the app would take this into account during the recruitment period. This limitation could be attributed to data collection via a link outside of the app platform because the engagement data showed higher numbers of people ($n=91$) continuing to use the app without completing the outcome measures. Owing to the anonymous nature of the evaluation, it was not possible to make comparisons between participants who downloaded the app and those who did not. Another limitation was the 2-week intervention period. Although this can be viewed as a strength for this preliminary evaluation, there may be benefits, such as observations about sustained use and further impact on well-being, if a longer app use period was included. Another limitation lies in the brief responses provided by the users for the semistructured questions. This limited the amount of data available to provide deeper insights into thematic analysis. These limitations and other key learnings will be considered in future evaluations of the app.

Future Directions

App Upgrades

Participants provided valuable feedback that had already been absorbed to inform upgrades to the app. On the basis of these findings, the following upgrades will be prioritized before further evaluations of the app are conducted. First, a proportion of participants expressed that they would like the app to be more interactive. A guided process has been added as an adjunct to the app so that users can join weekly sessions as part of a group to discuss sections of the app and reflect on concepts of self-awareness with an assigned coach. Second, the feedback indicated that some users forgot about the app and needed reminders. More in-app notifications have been included in newer versions of the app and are expected to further improve the levels of engagement [93,94]. Third, based on recent findings for increasing engagement [24,95] and the positive perception of the app's content and functionality, features such as personalized check-ins will be considered for inclusion in the app to further enhance the users' experience.

Research Implications

This evaluation provides a strong foundation for future evaluations of apps. The current recruitment, intervention, and data collection period (for a total of 6 weeks) may be too ambitious to gather a large enough sample size to meaningfully test statistical significance on multiple variables. On the basis of our findings and the experience of carrying out this evaluation, 12 weeks may be more feasible, allowing at least 1 month for each stage of the study. In this way, participants will

have a longer period to engage with the app and therefore complete all modules within the app. The outcome measures used appear to be appropriate, with almost 100% completion (for n=34 users) on all items. Future evaluations of the app could maintain the current data collection tools and include the anxiety measure (ie, GAD-7) at baseline to ensure that it accurately captures any direct influence of the app on participants' levels of anxiety. Therefore, this approach may go beyond the research question addressed in this evaluation and provide data on any relationship the app may have with lowering anxiety. This may prove to be valuable information that could substantiate the findings highlighted in the relationship between the app and depression scores. It may be just as useful to enhance the evaluation by adding a comparative group such as a blended intervention group and randomizing participants to ensure that our findings do not occur by chance. In future evaluations, it may be just as important for us to collect qualitative data from participants who drop out to fully capture if their decision to discontinue is based on app use or lack of time to complete outcome measures, as suggested by other researchers [96].

Practical Implications

Given the range of well-being apps, and the preliminary positive implications of this evaluation, improving well-being through regular use of an emotional well-being app that focuses on emotional awareness and knowledge of emotional patterns, could have potential promise as a low-cost approach to increase

well-being and related outcomes, such as life satisfaction. Sustaining user engagement in apps over time is not straightforward. The results of this evaluation are preliminary because of the moderate sample size across data collection points; however, there may be benefits to apps that draw on lived experience and storytelling, are multimodal, and promote increased emotional awareness and understanding of emotional patterns. This evaluation recruited a higher-than-average number of participants from minority ethnic backgrounds, and further apps and research initiatives ought to appeal to nonmajority app users.

Conclusions

This evaluation further highlighted the value of conducting formative research on mental health and well-being apps. For example, the findings suggest that users can provide valid feedback that can be used to inform future app upgrades. Users generally engaged with the app and provided favorable feedback regarding the app's content, functionality, and underlying principles. Notably, participants in this evaluation experienced significantly lower levels of depression scores after the intervention period, as well as increased well-being. The findings of this evaluation are encouraging and show positive preliminary evidence that suggests scope for further research with underrepresented groups, such as ethnic minority populations.

Acknowledgments

The authors thank the participants who took the time to participate in this study. The authors thank Rachel Gomes for supporting the study design and collecting objective app use data. The authors thank Isabella Vainieri, Fiby Saad, and Elisa Infanti for discussions in the early stages of the research and Nick Toner for his support when editing the early versions of the manuscript. This study was funded by Paradym, Ltd. Apart from commissioning a call for evidence in this area, CC had no role in data collection, data analysis, or data interpretation. However, CC was responsible for the oversight of the project and was involved in the study design and editing of the drafts of the manuscript.

Authors' Contributions

AE, CC, and SL conceptualized the study and contributed to the study design. All authors contributed to user testing. To reduce the risk of bias in favor of the intervention, SL and RMC independently analyzed the data and interpreted the output with input from AE and AMM. All authors contributed to the writing and editing of the manuscript. All authors were involved in reviewing, editing, and final approval of the manuscript before submission.

Conflicts of Interest

The authors declare no direct financial gain from conducting this study. However, it must be noted that CC is the founder and Chief Executive Officer of Paradym Ltd, and AE is employed as the research lead. To ensure researcher bias was not introduced, data analysis was conducted independently by at least 2 members of the research team. The other authors declare no conflicts of interest.

Multimedia Appendix 1

Ethical review decision tool developed by the Medical Research Council.

[[PDF File \(Adobe PDF File\), 170 KB - formative_v5i11e31064_app1.pdf](#)]

References

1. Agteren J, Iasiello M. Advancing our understanding of mental wellbeing and mental health: the call to embrace complexity over simplification. *Aus Psychol* 2020;55(4):307-316. [doi: [10.1111/ap.12440](https://doi.org/10.1111/ap.12440)]

2. Bell SL, Audrey S, Gunnell D, Cooper A, Campbell R. The relationship between physical activity, mental wellbeing and symptoms of mental health disorder in adolescents: a cohort study. *Int J Behav Nutr Phys Act* 2019 Dec 26;16(1):138 [FREE Full text] [doi: [10.1186/s12966-019-0901-7](https://doi.org/10.1186/s12966-019-0901-7)] [Medline: [31878935](https://pubmed.ncbi.nlm.nih.gov/31878935/)]
3. Huppert F. Psychological well-being: evidence regarding its causes and consequences. *Appl Psychol Heal Well Being* 2009 Jun 5;1(2):137-164. [doi: [10.1111/j.1758-0854.2009.01008.x](https://doi.org/10.1111/j.1758-0854.2009.01008.x)]
4. Chida Y, Steptoe A. Positive psychological well-being and mortality: a quantitative review of prospective observational studies. *Psychosom Med* 2008 Sep;70(7):741-756. [doi: [10.1097/PSY.0b013e31818105ba](https://doi.org/10.1097/PSY.0b013e31818105ba)] [Medline: [18725425](https://pubmed.ncbi.nlm.nih.gov/18725425/)]
5. Salari N, Hosseini-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Global Health* 2020 Jul 06;16(1):57 [FREE Full text] [doi: [10.1186/s12992-020-00589-w](https://doi.org/10.1186/s12992-020-00589-w)] [Medline: [32631403](https://pubmed.ncbi.nlm.nih.gov/32631403/)]
6. Kaufman K, Petkova E, Bhui K, Schulze T. A global needs assessment in times of a global crisis: world psychiatry response to the COVID-19 pandemic. *BJPsych Open* 2020 Apr 06;6(3):e48 [FREE Full text] [doi: [10.1192/bjo.2020.25](https://doi.org/10.1192/bjo.2020.25)] [Medline: [32250235](https://pubmed.ncbi.nlm.nih.gov/32250235/)]
7. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. *Sci Rep* 2021 May 13;11(1):10173 [FREE Full text] [doi: [10.1038/s41598-021-89700-8](https://doi.org/10.1038/s41598-021-89700-8)] [Medline: [33986414](https://pubmed.ncbi.nlm.nih.gov/33986414/)]
8. Kauer S. Emotional self-awareness and depressive symptoms: an investigation of an early intervention mobile phone self-monitoring program for adolescents. School of Psychological Science, The University of Melbourne. 2012. URL: <https://minerva-access.unimelb.edu.au/handle/11343/37899> [accessed 2021-10-04]
9. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016 Mar 01;3(1):e7 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
10. Keyes CLM. Mental health as a complete state: how the salutogenic perspective completes the picture. In: Bauer GF, Hämmig O, editors. *Bridging Occupational, Organizational and Public Health*. Dordrecht: Springer; 2014:179-192.
11. Agnoli S, Mancini G, Andrei F, Trombini E. The relationship between trait emotional intelligence, cognition, and emotional awareness: an interpretative model. *Front Psychol* 2019 Jul 30;10:1711 [FREE Full text] [doi: [10.3389/fpsyg.2019.01711](https://doi.org/10.3389/fpsyg.2019.01711)] [Medline: [31417455](https://pubmed.ncbi.nlm.nih.gov/31417455/)]
12. Frewen PA, Dozois DJ, Neufeld RW, Lanius RA. Meta-analysis of alexithymia in posttraumatic stress disorder. *J Trauma Stress* 2008 Apr;21(2):243-246. [doi: [10.1002/jts.20320](https://doi.org/10.1002/jts.20320)] [Medline: [18404647](https://pubmed.ncbi.nlm.nih.gov/18404647/)]
13. Hendryx MS, Haviland MG, Shaw DG. Dimensions of alexithymia and their relationships to anxiety and depression. *J Pers Assess* 1991 Apr;56(2):227-237. [doi: [10.1207/s15327752jpa5602_4](https://doi.org/10.1207/s15327752jpa5602_4)] [Medline: [2056418](https://pubmed.ncbi.nlm.nih.gov/2056418/)]
14. Guendelman S, Medeiros S, Rampes H. Mindfulness and emotion regulation: insights from neurobiological, psychological, and clinical studies. *Front Psychol* 2017 Mar 6;8:220 [FREE Full text] [doi: [10.3389/fpsyg.2017.00220](https://doi.org/10.3389/fpsyg.2017.00220)] [Medline: [28321194](https://pubmed.ncbi.nlm.nih.gov/28321194/)]
15. Chambers R, Gullone E, Allen NB. Mindful emotion regulation: an integrative review. *Clin Psychol Rev* 2009 Aug;29(6):560-572. [doi: [10.1016/j.cpr.2009.06.005](https://doi.org/10.1016/j.cpr.2009.06.005)] [Medline: [19632752](https://pubmed.ncbi.nlm.nih.gov/19632752/)]
16. Gross JJ, Muñoz RF. Emotion regulation and mental health. *Clin Psychol Sci Pract* 1995 Jun;2(2):151-164. [doi: [10.1111/j.1468-2850.1995.tb00036.x](https://doi.org/10.1111/j.1468-2850.1995.tb00036.x)]
17. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. *JMIR Mhealth Uhealth* 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: [10.2196/11926](https://doi.org/10.2196/11926)] [Medline: [31274112](https://pubmed.ncbi.nlm.nih.gov/31274112/)]
18. Roffarello AM, Russis LD. The race towards digital wellbeing. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow Scotland UK. [doi: [10.1145/3290605.3300616](https://doi.org/10.1145/3290605.3300616)]
19. Stawarz K, Preist C, Coyle D. Use of smartphone apps, social media, and web-based resources to support mental health and well-being: online survey. *JMIR Ment Health* 2019 Jul 12;6(7):e12546 [FREE Full text] [doi: [10.2196/12546](https://doi.org/10.2196/12546)] [Medline: [31301126](https://pubmed.ncbi.nlm.nih.gov/31301126/)]
20. Neary M, Schueller SM. State of the field of mental health apps. *Cogn Behav Pract* 2018 Nov;25(4):531-537 [FREE Full text] [doi: [10.1016/j.cbpra.2018.01.002](https://doi.org/10.1016/j.cbpra.2018.01.002)] [Medline: [33100810](https://pubmed.ncbi.nlm.nih.gov/33100810/)]
21. Eisenstadt A, Liverpool S, Infanti E, Ciuvat R, Carlsson C. Mobile apps that promote emotion regulation, positive mental health and well-being in the general population: systematic review and meta-analysis. *JMIR Ment Health* 2021 Jun (forthcoming) [FREE Full text] [doi: [10.2196/preprints.31170](https://doi.org/10.2196/preprints.31170)]
22. 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/)]
23. Flett JA, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness* 2018 Oct 31;10(5):863-876. [doi: [10.1007/s12671-018-1050-9](https://doi.org/10.1007/s12671-018-1050-9)]

24. Huberty J, Green J, Puzia M, Stecher C. Evaluation of mood check-in feature for participation in meditation mobile app users: retrospective longitudinal analysis. *JMIR Mhealth Uhealth* 2021 Apr 23;9(4):e27106 [FREE Full text] [doi: [10.2196/27106](https://doi.org/10.2196/27106)] [Medline: [33890865](https://pubmed.ncbi.nlm.nih.gov/33890865/)]
25. 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/)]
26. Bakker D, Rickard N. Engagement with a cognitive behavioural therapy mobile phone app predicts changes in mental health and wellbeing: MoodMission. *Aus Psychol* 2019;54(4):245-260. [doi: [10.1111/ap.12383](https://doi.org/10.1111/ap.12383)]
27. Birney A, Gunn R, Russell J, Ary DV. MoodHacker mobile web app with email for adults to self-manage mild-to-moderate depression: randomized controlled trial. *JMIR Mhealth Uhealth* 2016 Jan 26;4(1):e8 [FREE Full text] [doi: [10.2196/mhealth.4231](https://doi.org/10.2196/mhealth.4231)] [Medline: [26813737](https://pubmed.ncbi.nlm.nih.gov/26813737/)]
28. 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/)]
29. Throuvala MA, Griffiths MD, Rennoldson M, Kuss DJ. Mind over matter: testing the efficacy of an online randomized controlled trial to reduce distraction from smartphone use. *Int J Environ Res Public Health* 2020 Jul 05;17(13):4842 [FREE Full text] [doi: [10.3390/ijerph17134842](https://doi.org/10.3390/ijerph17134842)] [Medline: [32635650](https://pubmed.ncbi.nlm.nih.gov/32635650/)]
30. 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/)]
31. Ponso 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/)]
32. Litvin S, Saunders R, Maier MA, Lüttke S. Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: a randomized controlled trial. *PLoS One* 2020 Sep 2;15(9):e0237220 [FREE Full text] [doi: [10.1371/journal.pone.0237220](https://doi.org/10.1371/journal.pone.0237220)] [Medline: [32877425](https://pubmed.ncbi.nlm.nih.gov/32877425/)]
33. Daugherty DA, Runyan JD, Steenbergh TA, Fratzke BJ, Fry BN, Westra E. Smartphone delivery of a hope intervention: another way to flourish. *PLoS One* 2018 Jun 1;13(6):e0197930 [FREE Full text] [doi: [10.1371/journal.pone.0197930](https://doi.org/10.1371/journal.pone.0197930)] [Medline: [29856805](https://pubmed.ncbi.nlm.nih.gov/29856805/)]
34. Lindseth A, Norberg A. A phenomenological hermeneutical method for researching lived experience. *Scand J Caring Sci* 2004 Jun;18(2):145-153. [doi: [10.1111/j.1471-6712.2004.00258.x](https://doi.org/10.1111/j.1471-6712.2004.00258.x)] [Medline: [15147477](https://pubmed.ncbi.nlm.nih.gov/15147477/)]
35. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q* 1989 Sep;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
36. Dönmez-Turan A, Kir M. User anxiety as an external variable of technology acceptance model: a meta-analytic study. *Procedia Comput Sci* 2019;158:715-724. [doi: [10.1016/j.procs.2019.09.107](https://doi.org/10.1016/j.procs.2019.09.107)]
37. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008 Sep 29;337:a1655 [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
38. Bennett JA. The Consolidated Standards of Reporting Trials (CONSORT): guidelines for reporting randomized trials. *Nurs Res* 2005;54(2):128-132. [doi: [10.1097/00006199-200503000-00007](https://doi.org/10.1097/00006199-200503000-00007)] [Medline: [15778654](https://pubmed.ncbi.nlm.nih.gov/15778654/)]
39. Eisenstadt A, Liverpool S. User testing service evaluation. *Open Science Framework*. 2021 May. URL: https://osf.io/tvjba?mode=&view_only= [accessed 2021-10-13]
40. Hayes SC, Levin ME, Plumb-Villardaga J, Villatte JL, Pistorello J. Acceptance and commitment therapy and contextual behavioral science: examining the progress of a distinctive model of behavioral and cognitive therapy. *Behav Ther* 2013 Jun;44(2):180-198 [FREE Full text] [doi: [10.1016/j.beth.2009.08.002](https://doi.org/10.1016/j.beth.2009.08.002)] [Medline: [23611068](https://pubmed.ncbi.nlm.nih.gov/23611068/)]
41. Young J. *Cognitive Therapy for Personality Disorders: A Schema-Focused Approach*. Florida: Professional Resource Exchange; 1999.
42. Sutton A. Measuring the effects of self-awareness: construction of the self-awareness outcomes questionnaire. *Eur J Psychol* 2016 Nov 18;12(4):645-658 [FREE Full text] [doi: [10.5964/ejop.v12i4.1178](https://doi.org/10.5964/ejop.v12i4.1178)] [Medline: [27872672](https://pubmed.ncbi.nlm.nih.gov/27872672/)]
43. John OP, Gross JJ. Healthy and unhealthy emotion regulation: personality processes, individual differences, and life span development. *J Pers* 2004 Dec;72(6):1301-1333. [doi: [10.1111/j.1467-6494.2004.00298.x](https://doi.org/10.1111/j.1467-6494.2004.00298.x)] [Medline: [15509284](https://pubmed.ncbi.nlm.nih.gov/15509284/)]
44. Davis CH, Gaudiano BA, McHugh L, Levin ME. Integrating storytelling into the theory and practice of contextual behavioral science. *J Contextual Behav Sci* 2021 Apr;20:155-162. [doi: [10.1016/j.jcbs.2021.04.004](https://doi.org/10.1016/j.jcbs.2021.04.004)]
45. Mazzocco P, Green M, Sasota J, Jones N. This story is not for everyone: transportability and narrative persuasion. *Soc Psychol Pers Sci* 2010 Aug 06;1(4):361-368. [doi: [10.1177/1948550610376600](https://doi.org/10.1177/1948550610376600)]
46. Murphy S, Frank L, Moran M, Patnoe-Woodley P. Involved, transported, or emotional? Exploring the determinants of change in knowledge, attitudes, and behavior in entertainment-education. *J Commun* 2011 Jun 1;61(3):407-431. [doi: [10.1111/j.1460-2466.2011.01554.x](https://doi.org/10.1111/j.1460-2466.2011.01554.x)]
47. Bottorff JL, Sarbit G, Oliffe JL, Kelly MT, Lohan M, Stolp S, et al. "If I Were Nick": men's responses to an interactive video drama series to support smoking cessation. *J Med Internet Res* 2015 Aug 10;17(8):e190 [FREE Full text] [doi: [10.2196/jmir.4491](https://doi.org/10.2196/jmir.4491)] [Medline: [26265410](https://pubmed.ncbi.nlm.nih.gov/26265410/)]

48. Feldman G, Dunn E, Stemke C, Bell K, Greeson J. Mindfulness and rumination as predictors of persistence with a distress tolerance task. *Pers Individ Dif* 2014 Jan 01;56:10 [FREE Full text] [doi: [10.1016/j.paid.2013.08.040](https://doi.org/10.1016/j.paid.2013.08.040)] [Medline: [24298196](https://pubmed.ncbi.nlm.nih.gov/24298196/)]
49. Brinker JK, Chin ZH, Wilkinson R. Ruminative thinking style and the MMPI-2-RF. *Pers Individ Dif* 2014 Aug;66:102-105. [doi: [10.1016/j.paid.2014.03.001](https://doi.org/10.1016/j.paid.2014.03.001)]
50. Silva Almodovar A, Surve S, Axon DR, Cooper D, Nahata MC. Self-directed engagement with a mobile app (Sinaspriite) and its effects on confidence in coping skills, depression, and anxiety: retrospective longitudinal study. *JMIR Mhealth Uhealth* 2018 Mar 16;6(3):e64 [FREE Full text] [doi: [10.2196/mhealth.9612](https://doi.org/10.2196/mhealth.9612)] [Medline: [29549066](https://pubmed.ncbi.nlm.nih.gov/29549066/)]
51. 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/)]
52. Johansen KS. The use of wellbeing measures in primary health care/The DepCare Project. World Health Organization, Regional Office for Europe. 1998. URL: https://www.euro.who.int/_data/assets/pdf_file/0016/130750/E60246.pdf [accessed 2021-10-12]
53. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess* 1985;49:71-75. [doi: [10.1207/s15327752jpa4901_13](https://doi.org/10.1207/s15327752jpa4901_13)]
54. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. J Am Med Assoc* 1999 Nov 10;282(18):1737-1744. [doi: [10.1001/jama.282.18.1737](https://doi.org/10.1001/jama.282.18.1737)] [Medline: [10568646](https://pubmed.ncbi.nlm.nih.gov/10568646/)]
55. 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/)]
56. Brooke J. SUS: A 'Quick and Dirty' usability scale. In: *Usability Evaluation In Industry*. Florida: CRC Press; 1996.
57. Bech P. *Rating Scales for Psychopathology, Health Status and Quality of Life*. Berlin, Heidelberg: Springer; 2012.
58. Krieger T, Zimmermann J, Huffziger S, Ubl B, Diener C, Kuehner C, et al. Measuring depression with a well-being index: further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. *J Affect Disord* 2014 Mar;156:240-244. [doi: [10.1016/j.jad.2013.12.015](https://doi.org/10.1016/j.jad.2013.12.015)] [Medline: [24412323](https://pubmed.ncbi.nlm.nih.gov/24412323/)]
59. The satisfaction with life scale examining construct validity. *Statistics Netherlands*. 2012. URL: <http://C:/Users/user/Downloads/2012-09-x10-pub> [accessed 2021-10-04]
60. Vassar M. A note on the score reliability for the Satisfaction With Life Scale: an RG study. *Soc Indic Res* 2008;86(1):47-57. [doi: [10.1007/s11205-007-9113-7](https://doi.org/10.1007/s11205-007-9113-7)]
61. 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/)]
62. Kocalevent R, Hinze A, Brähler E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2013;35(5):551-555. [doi: [10.1016/j.genhosppsych.2013.04.006](https://doi.org/10.1016/j.genhosppsych.2013.04.006)] [Medline: [23664569](https://pubmed.ncbi.nlm.nih.gov/23664569/)]
63. Sousa TV, Viveiros V, Chai MV, Vicente FL, Jesus G, Carnot MJ, et al. Reliability and validity of the Portuguese version of the Generalized Anxiety Disorder (GAD-7) scale. *Health Qual Life Outcomes* 2015 Apr 25;13:50 [FREE Full text] [doi: [10.1186/s12955-015-0244-2](https://doi.org/10.1186/s12955-015-0244-2)] [Medline: [25908249](https://pubmed.ncbi.nlm.nih.gov/25908249/)]
64. Lewis J. The system usability scale: past, present, and future. *Int J Hum Comput Interact* 2018 Mar 30;34(7):577-590. [doi: [10.1080/10447318.2018.1455307](https://doi.org/10.1080/10447318.2018.1455307)]
65. Lewis JR, Brown J, Mayes DK. Psychometric evaluation of the EMO and the SUS in the context of a large-sample unmoderated usability study. *Int J Hum Comput Interact* 2015 Jul 31;31(8):545-553. [doi: [10.1080/10447318.2015.1064665](https://doi.org/10.1080/10447318.2015.1064665)]
66. Kern A, Hong V, Song J, Lipson S, Eisenberg D. Mental health apps in a college setting: openness, usage, and attitudes. *Mhealth* 2018 Jun 30;4:20 [FREE Full text] [doi: [10.21037/mhealth.2018.06.01](https://doi.org/10.21037/mhealth.2018.06.01)] [Medline: [30050916](https://pubmed.ncbi.nlm.nih.gov/30050916/)]
67. IBM SPSS Statistics for Windows, Version 27. IBM Corp. URL: <https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-27> [accessed 2021-10-04]
68. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
69. Koocher G, Keith-Spiegel P, editors. *Ethics in Psychology and the Mental Health Professions: Standards and Cases* (3rd Ed.). New York: Oxford University Press; 2008.
70. Ly KH, Dahl J, Carlbring P, Andersson G. Development and initial evaluation of a smartphone application based on acceptance and commitment therapy. *Springerplus* 2012 Jul 31;1:11 [FREE Full text] [doi: [10.1186/2193-1801-1-11](https://doi.org/10.1186/2193-1801-1-11)] [Medline: [23961342](https://pubmed.ncbi.nlm.nih.gov/23961342/)]
71. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res* 2010 Apr 30;12(2):e10 [FREE Full text] [doi: [10.2196/jmir.1371](https://doi.org/10.2196/jmir.1371)] [Medline: [20439251](https://pubmed.ncbi.nlm.nih.gov/20439251/)]
72. Ahtinen A, Mattila E, Väikkynen P, Kaipainen K, Vanhala T, Ermes M, et al. Mobile mental wellness training for stress management: feasibility and design implications based on a one-month field study. *JMIR Mhealth Uhealth* 2013 Jul 10;1(2):e11. [doi: [10.2196/mhealth.2596](https://doi.org/10.2196/mhealth.2596)] [Medline: [25100683](https://pubmed.ncbi.nlm.nih.gov/25100683/)]

73. Ly K, Ly A, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. *Internet Interv* 2017 Oct 10;10:39-46 [[FREE Full text](#)] [doi: [10.1016/j.invent.2017.10.002](https://doi.org/10.1016/j.invent.2017.10.002)] [Medline: [30135751](https://pubmed.ncbi.nlm.nih.gov/30135751/)]
74. Lui JH, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. *Profess Psychol Res Pract* 2017 Jun;48(3):199-210. [doi: [10.1037/pro0000122](https://doi.org/10.1037/pro0000122)]
75. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](https://pubmed.ncbi.nlm.nih.gov/27745683/)]
76. Liverpool S, Mota CP, Sales CM, Čuš A, Carletto S, Hancheva C, et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. *J Med Internet Res* 2020 Jun 23;22(6):e16317 [[FREE Full text](#)] [doi: [10.2196/16317](https://doi.org/10.2196/16317)] [Medline: [32442160](https://pubmed.ncbi.nlm.nih.gov/32442160/)]
77. Goodyear V, Armour K, Wood H. Young people learning about health: the role of apps and wearable devices. *Learn Media Technol* 2018 Nov 07;44(2):193-210. [doi: [10.1080/17439884.2019.1539011](https://doi.org/10.1080/17439884.2019.1539011)]
78. 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. [doi: [10.2196/20283](https://doi.org/10.2196/20283)] [Medline: [32990635](https://pubmed.ncbi.nlm.nih.gov/32990635/)]
79. Ha SW, Kim J. Designing a scalable, accessible, and effective mobile app based solution for common mental health problems. *Int J Human Comput Interac* 2020 Apr 26;36(14):1354-1367. [doi: [10.1080/10447318.2020.1750792](https://doi.org/10.1080/10447318.2020.1750792)]
80. Howells A, Ivtzan I, Eiroa-Orosa F. 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)]
81. Roy MJ, Costanzo ME, Highland KB, Olsen C, Clayborne D, Law W. An app a day keeps the doctor away: guided education and training via smartphones in subthreshold post traumatic stress disorder. *Cyberpsychol Behav Soc Netw* 2017 Aug;20(8):470-478. [doi: [10.1089/cyber.2017.0221](https://doi.org/10.1089/cyber.2017.0221)] [Medline: [28737954](https://pubmed.ncbi.nlm.nih.gov/28737954/)]
82. Deady M, Johnston D, Milne D, Glozier N, Peters D, Calvo R, et al. Preliminary effectiveness of a smartphone app to reduce depressive symptoms in the workplace: feasibility and acceptability study. *JMIR Mhealth Uhealth* 2018 Dec 04;6(12):e11661 [[FREE Full text](#)] [doi: [10.2196/11661](https://doi.org/10.2196/11661)] [Medline: [30514694](https://pubmed.ncbi.nlm.nih.gov/30514694/)]
83. Haeger J, Davis C, Levin M. Utilizing ACT daily as a self-guided app for clients waiting for services at a college counseling center: a pilot study. *J Am Coll Health* 2020 Jun 12:1-8 (forthcoming). [doi: [10.1080/07448481.2020.1763366](https://doi.org/10.1080/07448481.2020.1763366)] [Medline: [32529933](https://pubmed.ncbi.nlm.nih.gov/32529933/)]
84. 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/)]
85. Firestone R, Cheng S, Dalhousie S, Hughes E, Funaki T, Henry A, et al. Exploring Pasifika wellbeing: findings from a large cluster randomised controlled trial of a mobile health intervention programme. *N Z Med J* 2020 Oct 30;133(1524):82-101. [Medline: [33119572](https://pubmed.ncbi.nlm.nih.gov/33119572/)]
86. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. *NPJ Digit Med* 2019 Dec 2;2:118 [[FREE Full text](#)] [doi: [10.1038/s41746-019-0188-8](https://doi.org/10.1038/s41746-019-0188-8)] [Medline: [31815193](https://pubmed.ncbi.nlm.nih.gov/31815193/)]
87. Firth J, Torous J, Nicholas J, Carney R, Prapat A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017 Oct;16(3):287-298 [[FREE Full text](#)] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
88. Gál É, tefan S, Cristea IA. The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. *J Affect Disord* 2021 Jan 15;279:131-142. [doi: [10.1016/j.jad.2020.09.134](https://doi.org/10.1016/j.jad.2020.09.134)] [Medline: [33049431](https://pubmed.ncbi.nlm.nih.gov/33049431/)]
89. Eisenstadt M, Infanti E, Liverpool S. Characteristics and effectiveness of mobile apps that promote emotion regulation and well-being in the general population: a systematic review protocol. *National Institute for Health Research - PROSPERO* 2020. 2020. URL: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020213051 [accessed 2021-10-12]
90. Coelho CC, Tobo PR, Lacerda SS, Lima AH, Barrichello CR, Amaro E, et al. A new mental health mobile app for well-being and stress reduction in working women: randomized controlled trial. *J Med Internet Res* 2019 Nov 07;21(11):e14269 [[FREE Full text](#)] [doi: [10.2196/14269](https://doi.org/10.2196/14269)] [Medline: [31697244](https://pubmed.ncbi.nlm.nih.gov/31697244/)]
91. Walsh KM, Saab BJ, Farb NA. Effects of a mindfulness meditation app on subjective well-being: active randomized controlled trial and experience sampling study. *JMIR Ment Health* 2019 Jan 08;6(1):e10844 [[FREE Full text](#)] [doi: [10.2196/10844](https://doi.org/10.2196/10844)] [Medline: [30622094](https://pubmed.ncbi.nlm.nih.gov/30622094/)]
92. Weber S, Lorenz C, Hemmings N. Improving stress and positive mental health at work via an app-based intervention: a large-scale multi-center randomized control trial. *Front Psychol* 2019;10:2745 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2019.02745](https://doi.org/10.3389/fpsyg.2019.02745)] [Medline: [31866915](https://pubmed.ncbi.nlm.nih.gov/31866915/)]

93. Wasil AR, Venturo-Conerly KE, Shingleton RM, Weisz JR. A review of popular smartphone apps for depression and anxiety: assessing the inclusion of evidence-based content. *Behav Res Ther* 2019 Dec;123:103498. [doi: [10.1016/j.brat.2019.103498](https://doi.org/10.1016/j.brat.2019.103498)] [Medline: [31707224](https://pubmed.ncbi.nlm.nih.gov/31707224/)]
94. Wasil AR, Palermo E, Lorenzo-Luaces L, DeRubeis R. Is there an app for that? A review of popular mental health and wellness apps. *PsyArXiv* 2021:1-51. [doi: [10.31234/osf.io/su4ar](https://doi.org/10.31234/osf.io/su4ar)]
95. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. *JMIR Ment Health* 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: [10.2196/13717](https://doi.org/10.2196/13717)] [Medline: [31244479](https://pubmed.ncbi.nlm.nih.gov/31244479/)]
96. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2020 Feb 15;263:413-419. [doi: [10.1016/j.jad.2019.11.167](https://doi.org/10.1016/j.jad.2019.11.167)] [Medline: [31969272](https://pubmed.ncbi.nlm.nih.gov/31969272/)]

Abbreviations

ESAS-R: Emotional Self-Awareness Scale-Revised

GAD-7: Generalized Anxiety Disorder-7

PHQ-9: Patient Health Questionnaire-9

SUS: System Usability Scale

SWLS: Satisfaction With Life Scale

WHO-5: World Health Organization-Five Well-Being Index

Edited by G Eysenbach; submitted 19.08.21; peer-reviewed by J Huberty, LW Kranenburg, MM Abdul Kader Jilani; comments to author 13.09.21; revised version received 24.09.21; accepted 26.09.21; published 01.11.21.

Please cite as:

Eisenstadt A, Liverpool S, Metaxa AM, Ciuvat RM, Carlsson C

Acceptability, Engagement, and Exploratory Outcomes of an Emotional Well-being App: Mixed Methods Preliminary Evaluation and Descriptive Analysis

JMIR Form Res 2021;5(11):e31064

URL: <https://formative.jmir.org/2021/11/e31064>

doi: [10.2196/31064](https://doi.org/10.2196/31064)

PMID: [34569466](https://pubmed.ncbi.nlm.nih.gov/34569466/)

©Amelia Eisenstadt, Shaun Liverpool, Athina-Marina Metaxa, Roberta Maria Ciuvat, Courtney Carlsson. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 01.11.2021. 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 Social Determinants of Mental Health of Undergraduate Students in Bangladesh: Interview Study

Ananya Bhattacharjee¹, BSc; S M Taiabul Haque², PhD; Md Abdul Hady³, BSc; S M Raihanul Alam⁴, BSc; Mashfiqui Rabbi⁵, PhD; Muhammad Ashad Kabir⁶, PhD; Syed Ishtiaque Ahmed¹, PhD

¹Department of Computer Science, University of Toronto, Toronto, ON, Canada

²School of Computer Science and Mathematics, University of Central Missouri, Warrensburg, MO, United States

³Department of Computer Science and Engineering, Eastern University, Dhaka, Bangladesh

⁴Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

⁵Department of Statistics, Harvard University, Cambridge, MA, United States

⁶School of Computing, Mathematics and Engineering, Charles Sturt University, New South Wales, Australia

Corresponding Author:

Ananya Bhattacharjee, BSc
Department of Computer Science
University of Toronto
40 St George Street
Toronto, ON, M5S 2E4
Canada
Phone: 1 6476196982
Email: ananya@cs.toronto.edu

Abstract

Background: The undergraduate student population has been actively studied in digital mental health research. However, the existing literature primarily focuses on students from high-income nations, and undergraduates from limited-income nations remain understudied.

Objective: This study aims to identify the broader social determinants of mental health among undergraduate students in Bangladesh, a limited-income nation in South Asia; study the manifestation of these determinants in their day-to-day lives; and explore the feasibility of self-monitoring tools in helping them identify the specific factors or relationships that affect their mental health.

Methods: We conducted a 21-day study with 38 undergraduate students from 7 universities in Bangladesh. We conducted 2 semistructured interviews: one prestudy and one poststudy. During the 21-day study, participants used an Android app to self-report and self-monitor their mood after each phone conversation. The app prompted participants to report their mood after each phone conversation and provided graphs and charts so that the participants could independently review their mood and conversation patterns.

Results: Our results show that academics, family, job and economic condition, romantic relationship, and religion are the major social determinants of mental health among undergraduate students in Bangladesh. Our app helped the participants pinpoint the specific issues related to these factors, as the participants could review the pattern of their moods and emotions from past conversation history. Although our app does not provide any explicit recommendation, the participants took certain steps on their own to improve their mental health (eg, reduced the frequency of communication with certain persons).

Conclusions: Although some of the factors (eg, academics) were reported in previous studies conducted in the Global North, this paper sheds light on some new issues (eg, extended family problems and religion) that are specific to the context of the Global South. Overall, the findings from this study would provide better insights for researchers to design better solutions to help the younger population from this part of the world.

(*JMIR Form Res* 2021;5(11):e27114) doi:[10.2196/27114](https://doi.org/10.2196/27114)

KEYWORDS

Bangladesh; global south; social determinant; students; undergraduate; religion; women; mobile phone

Introduction

Similar to many low-and middle-income countries [1,2], mental health among youth in Bangladesh is an overlooked and stigmatized topic [3]. More than 30% of adults in urban areas struggle with mental health-related issues [4]. At the same time, mental health facilities and services in Bangladesh are also insufficient [5]. The National Institute of Mental Health and Pabna Mental Hospital are the only major institutes that provide mental health treatment, constituting only 700 beds in total [4]. There are some other mental health institutes in or around the big cities as well; however, two-thirds of the total population living in rural areas have difficulty accessing mental health because of the lack of facilities in rural areas [4]. Support for nonserious mental illness is even poorer; the number of mental and emotional health support helplines is extremely low [6]. In addition, the general population still holds a traditional negative attitude and stigma toward those with mental health issues [7,8].

Proposing a path forward for addressing mental health challenges in the Global South is still at a rudimentary stage of development [9]. The social determinants of mental health [10]—social and cultural factors that deeply impact one's mental health—have remained understudied in the context of Bangladesh. These factors vary depending on an individual's age, financial condition, or social surroundings. However, in this paper, we focus particularly on undergraduate students, who constitute a nontrivial part of the population and are susceptible to mental health-related problems [11]. As various statistics and news reports corroborate the deep-rooted existence of suicide [12], depression [13], substance use [14], and extremism [15] among the young population in Bangladesh, and previous research indicates that such behavioral aberrations are often fueled by broader social, cultural, and political contexts in which one grows up and lives in [16-21], it is timely to investigate these underlying factors.

Concerns over the mental health of university students have been expressed in the Global North for some time now [22], although the same cannot be said about the Global South. According to a study conducted in a European country, around one-third of the first-year university students have been found to have mental health-related problems [23]. Several studies [23-25] have associated academic performance with mental health; that is, poor academic performance worsens mental health and vice versa. Other contributing factors for mental health deterioration include economic problems, high parental expectations, strained relationships, and poor lifestyle [8,26-29]. Several mental disorders are frequently observed among people aged between 14 and 24 years [30], many of whom are also university students. These students tend to encounter new experiences such as moving away from home and making adult financial decisions. In addition, many of them experience changes in their health behaviors [31]. Adapting to these changes sometimes becomes difficult, as students are also expected to perform their coursework and participate in exams. Failure to cope with any of these challenges may cause mental health difficulties [32].

These factors have already been identified as the social determinants of the mental health of students [10,31], although they are often specific to the context of the Global North. However, the different structures of society and culture in the Global South [33] may contribute to people's mental health in other ways that are specific to the context of the Global South and have not been reported in previous studies. For example, in countries such as Bangladesh, extended families are far more prevalent where people live with their relatives (ie, uncles, aunts, cousins, or in-laws) [33]. Even when people are living in a nuclear family, social communication among relatives or neighbors is more frequent and cordial. As a result, maintaining regular communication with even distant relatives is viewed as an important social duty [34]. In addition, religion is a major component of people's lives [34]. The concepts of religion and religious duties differ significantly from those in the Western context. A large portion of the population finds happiness through practicing religious duties, as religion offers their lives a sense of value and purpose. Religious identity dictates an individual's social group and lifestyle choices [35]. The effects of complicated social relationships and disparate viewpoints on religion are yet to be understood in the context of students' mental health in the Global South.

Through this work, we intend to explore the social determinants of mental health, particularly in the context of Bangladesh. In a country like Bangladesh, where social and family structures are hierarchical, traditional, and community based [10,31,36], social and family relationships play an important role in people's daily lives, and recording one's mental condition after interactions with family members and social peers seems to be a useful way of monitoring and improving their mental health. Past research suggests that patterns of mobile phone calls can reveal important information about an individual's mental state and relationships with their friends and family [37]. In addition, mobile phone communication history helps one identify various forms of social closeness [38]. In recent years, Bangladesh has observed tremendous growth in mobile phone subscriptions. The total number of mobile phone subscribers in Bangladesh is approximately 160 million [39], which is almost equal to the total population of the country [40]. This shows that mobile phones are culturally relevant in Bangladesh. We see this as a research opportunity, as we can leverage the widespread use of mobile phones across the country to identify the social determinants of mental health of undergraduate students in Bangladesh.

One potential way toward forwarding the research on identifying the factors that affect mental health can be the use of reflective tools [41-43] that support self-monitoring of mood and emotions [44]. These mobile apps are suitable for measuring subtle emotions and capturing patterns of behavior in the long term [44]. A study conducted by Kauer et al [30] on youth aged between 14 and 24 years reported that self-monitoring one's mood results in an increase of emotional self-awareness, eventually reducing depression. Another similar study [45], which helped participants identify the cause of their drinking habits (eg, relation with romantic partners), showed success in tracking information about alcohol consumption and its effect on mental behavior. Kiekens et al [25] explored the impact of

mood-tracking apps in clinical settings, and their pilot study received positive feedback from both participants and therapists, as participants could review their relationships with their family and peers. Pollak et al [46] suggested that reflective tools with the option to provide self-reported emotions can give good insights into human behavior. In the psychology literature, ecological momentary assessment [47] supports the existence of a theoretical framework that demonstrates the importance of such in-the-moment mood assessment. The longitudinal nature of data in these tools gives users an opportunity to examine the effects of different life situations on their mental state, as users can document their moods and emotions at various times and do independent research on their historical data later [47]. Pollak et al [46] claimed that providing information about one's emotion and mental state at random times of the day can lead to improvement in social behavior. Other studies [48-52] have also shown the promise of self-reported mood-tracking apps in mental health research.

In this work, we designed, developed, and deployed a mood-tracking Android app that could be used as a reflective tool to record and reflect on past interactions with social peers. We conducted 2 semistructured interviews with our participants—one before and one another after using the app to understand the impact of our app on their mental health. Our primary goal is to progress the research toward identifying the factors that affect the mental health of Bangladeshi undergraduate students. However, we do not make any causal claims about these factors; instead, we present the findings from our interviews that give insights into how certain factors

manifest in the specific context of Bangladesh, as well as shed light on some new issues that were rarely discussed before. As a secondary goal of the study, we want to observe the effect of the self-monitoring tool and see whether our participants take any action to improve their mental health, despite no explicit recommendation provided by the app.

Methods

Participants

The study was conducted on 38 participants aged between 19 and 27 years. At the time of the study, all the participants were full-time undergraduate students. Participants were recruited through snowball sampling [53] via social acquaintances, classroom announcements, and word of mouth. Students were recruited from 7 universities in 2 major cities in Bangladesh: Bangladesh University of Engineering and Technology (BUET), Eastern University, Thengamara Mohila Sabuj Sangha Medical College, Dhaka College, Rajshahi University, City College, and Bangladesh University of Business and Technology.

Our participants included 30 traditional and 8 nontraditional students (P1, P5, P24, P25, P26, P30, and P35). In Bangladesh, traditional students join university immediately after receiving their high school diploma, whereas most nontraditional students attend university after receiving 2-3 years of technical vocational training (eg, application of different computer software in official contexts) and a few years of job experience [54]. [Table 1](#) provides information about the participants' gender, age, and institution.

Table 1. Information about the participants^a.

Participant	Gender	Age (years)	Institution	Student type
P1	Male	25	EU ^b	Nontraditional
P2	Male	24	EU	Traditional
P3	Male	22	EU	Traditional
P4	Female	24	EU	Traditional
P5	Male	25	EU	Nontraditional
P6	Female	24	EU	Traditional
P7	Female	25	EU	Traditional
P8	Male	23	BUET ^c	Traditional
P9	Male	23	BUET	Traditional
P10	Male	24	BUET	Traditional
P11	Male	23	BUET	Traditional
P12	Male	23	BUET	Traditional
P13	Male	23	BUET	Traditional
P14	Male	24	BUET	Traditional
P15	Female	23	BUET	Traditional
P16	Female	23	BUET	Traditional
P17	Female	23	BUET	Traditional
P18	Female	20	BUET	Traditional
P19	Male	23	BUET	Traditional
P20	Female	22	RU ^d	Traditional
P21	Male	22	DC ^e	Traditional
P22	Male	21	DC	Traditional
P23	Male	22	DC	Traditional
P24	Female	24	EU	Nontraditional
P25	Male	27	EU	Nontraditional
P26	Male	20	EU	Nontraditional
P27	Female	22	CC ^f	Traditional
P28	Male	19	RU	Traditional
P29	Male	20	RU	Traditional
P30	Male	25	BUBT ^g	Nontraditional
P31	Female	19	TMC ^h	Traditional
P32	Female	19	TMC	Traditional
P33	Female	19	TMC	Traditional
P34	Female	19	TMC	Traditional
P35	Male	26	EU	Nontraditional
P36	Female	23	EU	Traditional
P37	Male	23	BUET	Traditional
P38	Female	20	TMC	Traditional

^aP36-P38 did not take part in exit interviews.

^bEU: Eastern University.

^cBUET: Bangladesh University of Engineering and Technology.

^dRU: Rajshahi University.

^eDC: Dhaka College.

^fCC: City College.

^gBUBT: Bangladesh University of Business and Technology.

^hTMC: Thengamara Mohila Sabuj Sangha Medical College.

Study Design

Overview

We designed and developed *HWC* (How Was the Call?)—a mobile phone app that is a self-reporting tool to record and reflect on past phone calls. We conducted 2 semistructured interviews with each participant, one before (baseline interview) and another after (exit interview) using the app. Between these 2 interviews, the participants were asked to install and use the app for 3 weeks. The study was approved by the institutional review board of the authors' institution.

Baseline Interview

During this interview, we asked the participants about their general mental health and the reasons behind their mental stress, depression, or frustration; the frequency of experiencing mental health issues; the relationships that cause them; and their coping strategies. As English is the primary medium of instruction in their institutions, all of our participants were proficient in English. However, to collect spontaneous responses and have an in-depth discussion, we conducted the interviews in their native language Bengali, the official language of Bangladesh [55]. We occasionally used English words for different mobile phone and mental health–related terms.

Interviews were conducted both in person and through the Zoom videoconferencing platform. All interviews were audio recorded with the consent of the interviewees. Interviewers ensured appropriate levels of empathy and rapport during the interviews. Time-outs were offered if the participants asked for them.

The average length of the interviews was 25 minutes (SD 13.76; SE of the mean 3.97).

The 3-Week Study With the HWC App

After the baseline interview, participants were requested to use the HWC app for 3 weeks. Study personnel helped the participants install the app on their smartphones and instructed them on how to use it. We designed the app to run in the background and pop up a survey at the end of each phone call. The survey asked the participants to describe their mood by choosing from 8 emotions (Figure 1): *excited, cheerful, calm, neutral, bored, sad, tense, and irritated*. These 8 emotions cover the circular space of the valence-arousal dimension [56], a common framework for recording emotional experience, which assumes that all human emotions are distributed in a 2D space. Previous works [37,44,46] show that this approach of asking emotions from the valence-arousal dimension is particularly suitable for apps where users need to give frequent quick inputs on their mood, and it offers a reliable tool for researchers to analyze user behavior. One might argue that self-reported moods might not cover all mood archetypes; however, a vast amount of psychology literature [37,44,46] validates the use of our or similar to our self-report methods. In addition to the ease and convenience of the implementation of these methods, self-reported methods provide users with enough control over their interaction with the system, which does not make them feel they are part of an experiment and instead generate data closer to the real world [44,57]. These approaches inspire spontaneous responses from people and are helpful for capturing fluctuations in behavior and symptoms [57].

Figure 1. Interface for providing feedback after each call. HWC: How Was the Call?

HWC!

How are you feeling right now after communicating with this person?

Name

Excited Cheerful Calm Neutral

Bored Tensed Sad Irritated

Additional comments about the call

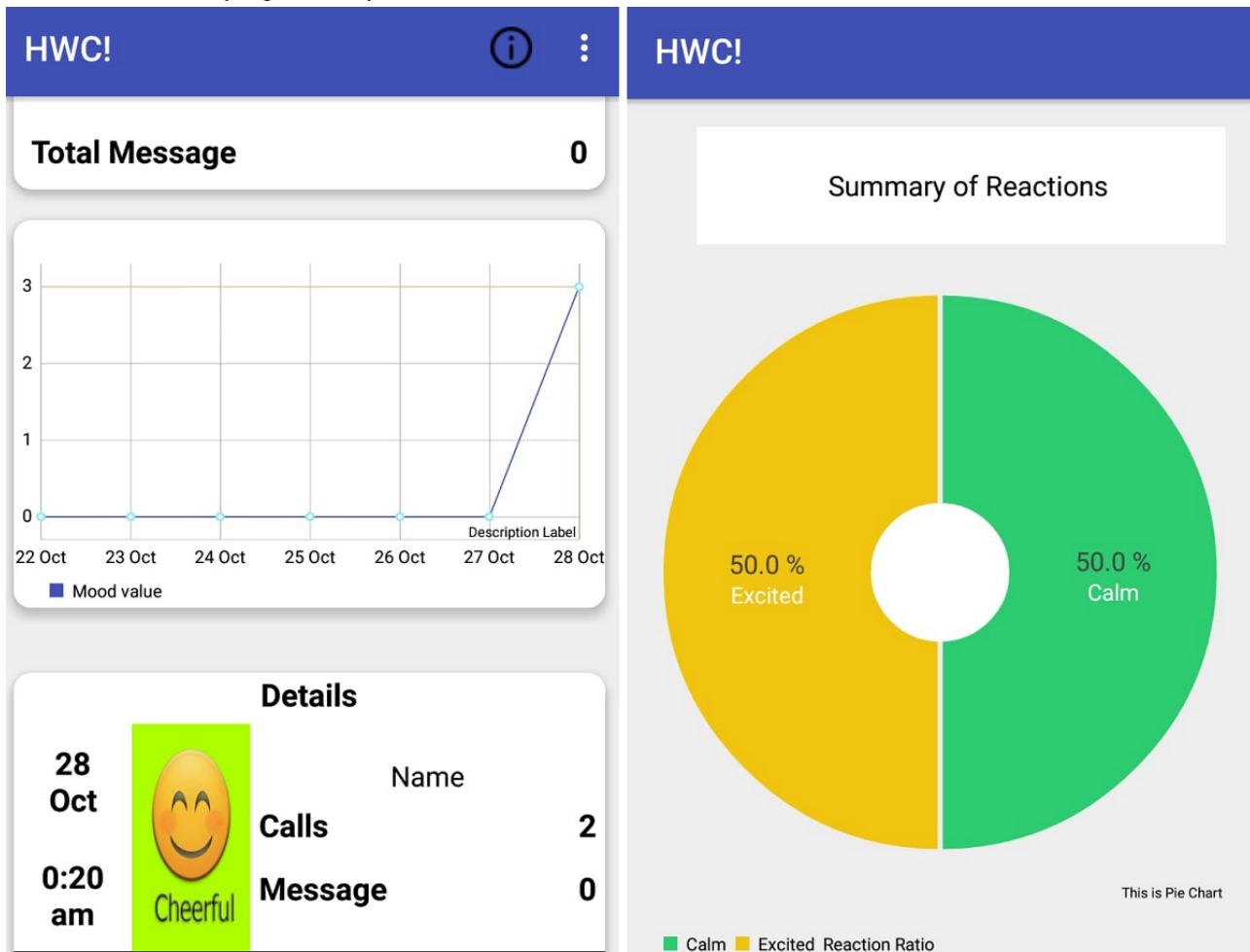
SKIP NEXT

In addition to the 8 emotions, we also provided the participants with a textbox to describe their feelings in more detail. Although we could have asked for more inputs from users (eg, how strong is the emotion?), our app asked for user input every time they made or received a call. Hence, asking for too many inputs from users after each phone call would require a high cognitive burden that might alter emotion and reduce long-term engagement [44]. However, a participant can choose to skip and provide no information. This app does not require any internet connection and stores all information in the local

memory of the mobile phone. As a result, there is no risk of participants' information going public.

Once the participants described their feelings, they had the option to view their data later. The participants could review their call history and see how they felt after each call (Figure 2). They could also view a visual summary of their reactions over a period of time with the help of a pie chart (Figure 2). Although the app let participants see their information, it did not provide any sort of feedback or recommendation to alter the mood.

Figure 2. Interfaces for analyzing call history and mood. HWC: How Was the Call?



The flow of the app is as follows:

1. The participant installs the app.
2. Every time the participant makes or receives a regular phone call, the app starts its operation.
3. After the call ends, the participant is prompted to provide feedback and choose from 8 options (Figure 1). A textbox is provided for additional written comments. A participant can take any of the following 2 steps:
 - After providing the information, the participant can press Next. This stores the information in the local memory of the mobile phone, and the app terminates.
 - The participant can choose not to provide any information by pressing Skip. The app then terminates.
4. The participant checks the call history and reviews their emotional status after each conversation at any time (Figure 2).

Participants who were worried about their data being stored outside the phone (public server and database) or shared with a third party were ensured that their call history and mood descriptions would only be stored in their mobile phone memory, and they were even allowed to see the source code of the app. In fact, 2 participants who were familiar with Android app development requested to review our source code, and we allowed them to do so. We did not collect specific information

about the number of calls received by each participant for privacy reasons.

Exit Interviews

Exit interviews were arranged after the participants had used the HWC app for 3 weeks. In these interviews, participants described their experience of using the app to track their moods. The questions were more specific in this session, where we asked them to highlight the major incidents that they had recorded over the past 3 weeks. We also asked them how the app had shaped their opinions and attitudes toward other people in their social circle. The setting of this interview was the same as that of the baseline interview. The average length of the interviews was 17 (SD 7.43, SE 3.03) minutes. Despite our best efforts, we could not reach 3 participants (P36-P38) in this session because of their unavailability, resulting in a data set of 35 (male: 21/35, 60%; female: 14/35, 40%) interviewees.

Data Analysis

We first transcribed the audio recordings from both sets of interviews and then translated them into English. Next, we performed a thematic analysis of our data based on the Boyatzis framework for code development [58]. From our analysis of the baseline interviews, we tried to explore the major social determinants of mental health among undergraduate students in Bangladesh. The responses of the participants during the exit interviews demonstrated the effectiveness of the app in

monitoring their mental health and its impact on identifying the specific personal issues they had been dealing with. The results of the analysis are described in the next section.

Results

Overview

In this section, we first describe the findings of our baseline interviews by highlighting the factors that affect the mental health of our participants. Then, we report the impact of our app on helping participants monitor their mental health. Finally,

we describe the measures that were taken by our participants to improve their mental health, although no recommendation was provided by the app.

Factors Affecting Mental Health

Overview

Through our interviews, we identified 5 factors that affected our participants' mental health the most. Table 2 illustrates the distribution of these factors based on how many participants reported them. Several participants reported multiple factors. Each of these factors is presented in detail later.

Table 2. Distribution of the factors that were reported by the study participants (n=35).

Factor	Participants, n (%)
Academic performance	27 (77)
Family	19 (54)
Job and economic condition	17 (49)
Religion	11 (31)
Romantic relationships	9 (26)

Academics

Academic standing was the most discussed theme during the interviews. Of the 35 participants, 27 (77%) noted that academic performance had a major impact on their mental health. Their main concern was the stress and depression created by assignments, class tests, projects, and final exams. One participant said:

I feel depressed on the evening before my exams. I get very tensed and start to think too much about the upcoming exam. Often I feel uncertain whether I will be able to answer all the questions or not. Sometimes I am too confused and start to panic about my preparation. I call my classmates and ask them about their preparation. Although they tell me that they have not prepared very well either, I assume otherwise and get more depressed. [P4]

Another participant (P11) further added that he had been stressed because of his final-year academic thesis. Even during vacation, he had to work on his thesis. Other participants were more concerned about their academic grades. They thought that their results were not satisfactory, and they might struggle in the job market. As mentioned by another participant:

I get frustrated when I try to read in the evening. I think about my future career and feel that I am not doing enough. I have a poor CGPA, will anyone give me a job? At times, it seems that I am a complete failure, which makes me frustrated. [P2]

A similar sentiment was reflected in the response of P10. In his opinion, despite putting in his best efforts, he had not been able to achieve his desired grades, which resulted in mental depression and frustration. He said the following:

I have been getting poor grades in exams in consecutive semesters. Despite trying my best, I am

not being able to improve my grades. It's not that I'm not trying, but nothing seems to work. [P10]

Another participant (P17) expressed her concerns over academic deadlines. She felt that procrastination contributed to a decline in her mental health. She tended to start working at the 11th hour to complete her assignments; however, before that, she could not focus on anything else either, as she constantly kept reminding herself about the deadline. She felt that this sort of behavior put her in *some kind of loop*, which was the main reason behind her stress. P25 also shared similar frustrations; however, his reason was that he did not have time to focus on his studies because of his full-time job. He also felt that he was in a loop as he joined the program to have a BSc degree and have promotions in his current career; however, at the same time, his career prevented him from concentrating on his studies.

Family

Of the 35 participants, 19 (54%) mentioned family as a reason for the decline in their mental health. However, the underlying reasons varied significantly. As the participants were from different backgrounds in terms of social and economic status, their problems were also different in nature. For example, one participant (P1) said that he had lost his parents at the age of 15 years. As the eldest child, he had to pick up many responsibilities since then. Although he was quite happy to perform his responsibilities, he constantly felt that he had not been doing enough for a better upbringing of his younger brother. He said the following:

When my mother died, my younger brother was still a baby. We lost our father after a few years, so his only guardian now is me. I am always tense about his education. Sometimes my brother does not listen to what I say. Those times are very hard for me. Whenever he fails at anything, I feel genuinely frustrated. [P1]

The participants who were married (P5, P6, P7, and P24) talked about their spouses and in-laws when discussing mental issues. One male student (P5) with a full-time job said that he was always concerned about financial stability, as his wife did not have a job. Owing to his busy life, he missed many family events, and although the reasons were genuine, his in-laws felt that he did not give them enough attention. Of the 35 participants, the 3 (9%) recently married female participants (P6, P7, and P24) discussed at length about their relationships with in-laws and how this affected their mental health. In Bangladesh, it is a common practice for women to move to their husbands' houses after the wedding. As extended families are still a cultural norm in Bangladesh [33], a new wife needs to build a relationship with her husband's family. These issues were pointed out by both P6 and P7. P7 said the following:

I am currently having issues with my new family. It is not any particular person, but their overall attitude. I have to be careful about my every action because if they don't like anything, they directly call my parents to complain about me. What hurts me more is that my own parents do not support me in this case. The relationship dynamic with my in-laws is perhaps the biggest cause of stress in my life currently. [P7]

Some other participants mentioned that as they were too busy with their academic life, they got stressed whenever they were burdened with any additional family responsibilities, including sending money and taking care of sick family members. Approximately 6% (2/35) of participants (P13 and P27) said that simply talking about family problems through the mobile phone was stressful for them.

Job and Economic Condition

Of the 35 participants, 5 (14%) participants (P1, P5, P25, P30, and P35—all nontraditional students) had full-time jobs during the study period, and all remarked that work-related issues affected their mental health. P5 talked about the competitive environment of his office—how everyone there was under huge pressure to perform and complete the tasks assigned to them. He said the following:

A full-time employee has a lot of responsibilities. You have pressure from your boss – pressure of meeting deadlines, pressure of performing better than your colleagues. Even at midnight, I receive calls from my boss. This makes me think that life without mobile phones would have been much better. At least your boss could not have reached you after office hours. [...] When I fail to meet a deadline, I get anxious. You fail, which means that you will have a lesser chance of being promoted and there are many other people waiting to grab that opportunity. I used to think competition exists only in colleges, but it is certainly more in professional life. [P5]

Some of our participants mentioned that their families were solvent, and they did not have an urgency to get internships or jobs. However, they still felt the pressure to look for a job, as their peers had been doing the same. One participant (P14) highlighted this issue as a major reason for the decline in his mental health. Some participants, on the other hand, got

confused about their career choices. In Bangladesh, career counseling is not recognized as an important activity at universities [59,60]; thus, many students face a dilemma regarding their career choices. This is reflected in the response of a business major:

I think about my future all the time. I am finding it really difficult to choose between two specific options: doing an MBA or preparing for job interviews. I try to consult my seniors, but even they do not have a clear career path for this program. [P6]

As mentioned before, of the 35 participants, 12 (34%) were from BUET, which is typically regarded as the most sought-after academic institution in the country. As a result, they had a high demand as instructors in local coaching centers [61] or as private tutors. In fact, all of our participants from BUET, except P9, had a part-time income source as private tutors. We found a few tutoring-related issues that affected their mental health. For example, P10 mentioned that he received many calls from his students during his own study hours, and concentrating too much on his students might be a reason for the deterioration of his mental health as it affected his own grades. P11 also expressed similar concerns; however, for him, making money was a bigger priority than obtaining good grades, even as a student. In his opinion, he became frustrated when he did not have a good number of students to be able to send enough money to his family. In fact, he was mentally quite happy during our study period as he had been earning more money than ever.

Religion

Of the 35 participants, 11 (31%; all Muslims) discussed the role of religion in their mental health. One of them (P1) said that when he missed the Fajr prayer (the first of the 5 daily prayers performed by a practicing Muslim) at dawn for oversleeping, he did not feel well for the rest of the day. However, he eventually calmed himself by thinking that an unwillingly committed negligence would not be punished and a good act later in the day would compensate for the earlier remission. He further added the following:

I believe in Tawhid [the concept of monotheism in Islam], so I believe the whole universe has been created by our God. He has given us certain duties to perform, and sometimes when I fail to do any of those, I get a little bit stressed out. [...] I have another issue pertinent to religion. In a society, different types of religions coexist and I have certain expectations from the believers of other religions. At times, those expectations are not met and I face trouble in interacting with people from other religions, which bothers me a lot mentally. [P1]

P27, who also expressed similar sentiments, felt that each problem in life could not be shared with family. That is why he had turned to God with Whom he feels that every problem can be shared and discussed. However, all of these participants highlighted that they performed Salah (prayer) to seek happiness. In addition, P4 and P24 said that they regularly read the Quran (the main religious text of Islam) to relieve mental stress.

Romantic Relationships

Of the 35 participants, 9 (26%) talked about their romantic relationships when discussing mental health. The 3 married participants, including P7, who was not happy with her in-laws, appreciated the support they received from their spouse in conjugal life. However, all of them confessed that they needed to think about their actions thoroughly, as those might affect their spouse. This constant pressure of remaining careful and alert creates a stressful family environment for them.

The unmarried participants (P9, P11, P13, P21, P28, and P31) described their past and present romantic relationships. One participant correlated his stress and frustration with his relationship with his partner:

I think my major source of stress and depression is my relationship with my girlfriend. We have been in the relationship for a while now, and I always try to make her happy in any way I can. Even then, I sometimes notice that the value of my opinion matters less very little in our relationship. Additionally, I have a tendency to compare my financial status with my girlfriend's, which also makes me sad or in some cases, jealous. [P21]

Another participant (P9) pointed out a different aspect of relationships, saying that his ex-girlfriend had issues of emotional dependency and that he had to spend a significant amount of time throughout the day talking with her over the phone. He felt relieved to get out of the relationship as he was unable to cope with the habit of spending too much time on the phone. P11, who had recently broken up with his girlfriend, was visibly distressed because of the failed relationship. However, one participant (P13) attributed the supportive nature of his girlfriend to his sound mental health, saying that he always found his partner by his side when he was stressed.

Monitoring Mental Health

Having described the factors affecting the mental health of our participants, we now highlight the effectiveness of our app in monitoring their mental health. Through our app, participants were able to track their own behavioral patterns by regularly recording their information. All of our participants agreed that the moods shown in the app represented the entire spectrum of their real-life emotions and made them more conscious about their mental health:

Each time I finish a call, the app asks me how I am feeling. This makes me think about my mental health for a few moments, which I would not have done previously. The app makes me more aware of my mental health. [P1]

Another participant (P2) liked the feature of reviewing the mood data. He reported that during the course of 3 weeks, he had periodically analyzed his mood report, and in doing so, he could identify some patterns. For example, he had realized that he felt more tense at night.

Our participants also mentioned that the app helped them in identifying the positive and negative influences in their lives. As a result of using the app, they were able to more consciously

identify the positive influencers—persons with whom they are close in real life but never previously thought of as positive catalysts for the betterment of their mental health. For example, P5 acknowledged this revelation in the following way:

Through this app, I could identify those people who make me happy and I realized that even during face-to-face conversations, this same group of people makes me feel more comfortable. I could make this explicit connection by using this app. [P5]

Similarly, some participants noted that the use of the app was helpful in identifying persons who caused mental stress in their lives. Of the 35 participants, 2 (6%) participants (P6 and P26) mentioned that the emojis were quite helpful in this regard as they could properly categorize their mental state by using those emojis. They were able to identify the conversations that had a negative impact by looking at the emojis for *sad* and *irritated* reactions. A couple of other participants (P2 and P4) also explicitly recognized the importance of the emojis. Choosing emojis after each phone call gave P4 a sense of excitement similar to *playing games on a mobile phone*, which she viewed positively.

Some of our participants reported how the app assisted them in precisely identifying the reason for their mental stress. For example, during the baseline interview, P11 said that the final-year thesis had been his primary academic concern. By using the app, he was able to understand that his supervisor had been pushing him a lot for the thesis. After each phone call with his supervisor, he was either *tense* or *sad*. At times, he felt *irritated* after receiving several calls regarding the improvement of the thesis document. Similarly, before using the app, P13 already knew that talking with his family members over the phone had been stressful for him. After using the app for 3 weeks, he realized that things got worse when economic problems were discussed during any conversation but at other times, the conversations were not as stressful. Of the 35 participants, 7 (20%) participants (P3, P10, P18, P22, P28, P30, and P35) specifically mentioned that the app helped them pinpoint the exact cause of their mental stress.

One participant (P12) provided important feedback regarding the textbox feature for writing additional comments. He thought that the overall mood was not entirely dependent on the conversation, as it got affected by other factors (eg, weather or an upcoming exam). He particularly liked the option of having a textbox to be able to record the potential reasons behind certain emotions. In his opinion, the written comments were helpful to better understand the contexts when he reviewed them later.

Impact on Mental Health Improvement

As the use of the app enabled our participants to better identify the persons or events that had a negative impact on their mental health, they could also take some measures for improvement. About two-thirds (24/35, 69%) of our participants said that they had tried to change the frequency of contacting those people who were having a negative effect on their mental health:

The app helped me identify the people who were causing stress. As I checked the call patterns, I tried to understand why I was not feeling comfortable

talking with them. What I did was – I tried to improve my relationship with them. Some attempts succeeded, while others did not. I have started avoiding those persons with whom I could not develop a better relationship. [P6]

Responses from P14 also revealed similar adjustments. By using the app, he realized that he had been in touch with some friends who were detrimental to his mental health as phone conversations with them were making him *sad*. He reduced the frequency of phone conversations with those people and reported that he had already started feeling better. P29 was also able to identify such people in his life. Apart from having fewer phone conversations with them, he started changing his tone and approach in physical meetings:

After I found out who those friends were, whenever I met them I explicitly tried to show that I am not interested in talking with them. I deliberately made the conversations shorter and removed the friendly approach from my tone. [P29]

However, our participants also mentioned that it is impossible to avoid certain persons in academic or professional lives (eg, supervisors and managers), and they had to pick up phone calls from those persons or call them for important academic or professional reasons. At times, they also had to pick up specific calls for courtesy and modesty. Some participants mentioned that the app had been helpful in these cases as based on previous mood patterns recorded through the app; they were mentally prepared for certain conversations to have a potential negative impact on their mental condition:

In my office, I often need to contact many colleagues and those conversations can't be avoided. But now the app has let me know that after some of those conversations, I'll not feel cheerful. As I know this beforehand, the impact is less severe. [P5]

One participant (P1) said that the app helped him control his temper. He talked about a particular relative who used to call at inappropriate times for mundane conversations. As P1 felt *annoyed* after receiving that relative's call, he tried his best to remain calm when talking with him. P1 also mentioned that the app made him aware of his attitude toward a few other family members as well. Another participant (P10) said that the app offered him the opportunity to find a new support system within his own family:

I have a large family with many sisters. As I am under huge academic pressure, I cannot contact them as frequently as I would like to. But after using the app, I found that my stress relieves a lot when I share my problems with my sisters. They give wonderful advice, too. I have decided to communicate more frequently with them. [P10]

The app also seemed to have a positive impact on the relationship status of a participant (P13). Although P13 had a stable and healthy relationship and used to share his problems with his girlfriend even before using the app, he noted that the app reinforced the same notion about the relationship, as he often felt *cheerful* or *calm* after talking with her.

Discussion

Principal Findings

Social Determinants of Mental Health of Undergraduate Students in Bangladesh

Our baseline interviews identified several factors that affect the mental health of students in Bangladesh. Our findings demonstrate some differences between Western countries and Global South countries such as Bangladesh with regard to social determinants of mental health of the young, college-going population. Although some of the factors (eg, academic performance [23-25,62] and economic conditions [28,63]) we identified in our study have already been pointed out by previous studies conducted in Western contexts, issues such as extended family (in-laws) problems and religion have not been reported before.

Female students who go for higher education need to overcome the barriers created by the local society [64,65]. This situation is even more difficult for married students. In a country where >52% of children get married before reaching 18 years of age [66,67], it is expected that a large number of undergraduate students, especially females, would be married. We had 3 married students in our study, 2 of whom were female. According to the local culture of Bangladesh, the bride moves in with the groom's family after marriage. New brides often need to adjust to a different family and adapt to their culture and rules. When discussing mental health, both of our female married participants talked about their struggle to adapt themselves to their husbands' families. This struggle to adapt with in-laws or extended families should be investigated in the context of students from other countries in nearby regions (eg, India and Pakistan), as underage marriage and discrimination against women likely exist in those countries as well [2,68].

Religion has been an integral part of the social, political, and economic environments of Bangladesh [34]. The state religion of the country is Islam [69], and 90.3% of the total population identifies themselves as Muslim [40]. Religious identity is a significant factor behind happiness in Bangladesh [34]. This view aligns with one of our participants' stress behind the failure to perform religious duties. Some other participants also briefly mentioned that they performed regular prayers to seek mental happiness.

Although economic conditions have been a factor discussed in previous studies, the context of Bangladesh is somewhat different. In the local society, a male is expected to support his parents and, in some cases, the extended family. Therefore, students at the undergraduate level, especially male students, start worrying about their financial status. Families also start sharing financial problems with them, indicating that the students now need to contribute to their families. For example, P11 and P13 felt stressed and frustrated when they failed to send enough money to their homes. However, we did not get such reports of stress over the family's economic status from any of our female participants.

Apart from traditional students, we also had 23% (8/35) nontraditional students (female: 1/8, 13%; male: 7/8, 87%) in

our study. Like regular students, they mentioned issues about coursework and exams, but all of them pointed out other issues as well. Of the 8 nontraditional students, 4 (50%) were married; therefore, these participants mentioned their struggle to measure up to their in-laws. Even the male married student (P5) talked about how his inability to attend family functions made his in-laws unhappy. Of the 35 participants, 5 (14%) male participants (P1, P5, P25, P30, and P35) explained their job life in detail and expressed that they found it very difficult to manage a full-time job with academic life. Their conversations (particularly that of P25) hinted that they were expected to do a full-time job as they were getting old; however, at the same time, they had to be serious about their academic coursework, as not having a bachelor's degree hindered their progress in job life.

How the App Helps Identify One's Own Social Determinants

The most commonly appreciated theme regarding our app was its ability to help users observe their mental states. In the baseline interviews, many of our participants had rough ideas about the main factors affecting their mental health; however, using the app for 3 weeks helped them pinpoint the exact reason. During this period, our participants constantly put their emotions after each call. As our app provided visual analysis of call history through graphs and pie charts, they could interpret those in their own way. Our participants were able to notice which calls evoked negative sentiments and acted accordingly. For example, the app helped several participants identify that academic supervisors or some specific family members were the reason for their stress. Conversely, the app also helped them identify the relationships that had a positive impact on their mental health. One participant further discovered the temporal aspect of mental health by using the app.

Similar to other reflective tools [41-43] with the option of providing self-reported emotions, a benefit of our app was that the participants could analyze their calls and moods in any way they wanted. Looking at the graphs and charts, they could review their moods over multiple days, or they had the option to analyze their conversations with a single individual. Although the participants mostly talked about identifying positive or negative relationships in our study, we also had people who identified other patterns of their mental states (eg, P2 felt more stressed at night). P13 identified that talking with family only stressed him when the conversation was about financial problems. Overall, the flexibility that we provided to our participants regarding the interpretation of their calls and emotions enabled us to have a diverse set of insights into their mental state.

Although the app was used only by Bangladeshi students in our study, apps such as these can be used in a universal context. In our study, students could identify the factors causing a negative impact on their mental health. These factors may be different in the context of people from other cultures; however, everyone can realize their own factors by using the app for a prolonged period. Even in our study, the participants reported a range of factors using the same app. We believe that this app can be considered as a variant of a digital diary [70], which can act as a support system for mental health [11]. However, we note that

the topic of using the textbox to describe the phone calls had come up only once in our study.

Actions Taken by Users

After identifying some of the potential factors, the participants in our study took some measures to improve their mental health. These measures were not suggested or recommended by our app; rather, they were taken by the participants on their own. For example, 24 students reported that they changed the frequency of contacting the people whom they found to have a negative impact on their lives. Among these participants, some tried adjusting their stressful relationships, whereas others started avoiding interactions altogether. In cases where avoiding a person was not a feasible option because of professional reasons (eg, colleagues or supervisors), the app helped the participants prepare themselves mentally for a potential negative conversation.

It should also be noted that the app cannot address all the cultural or social aspects of human life. For example, our mobile app cannot directly assist female students regarding the challenges they face in their day-to-day lives. Treating women as inferior or violence against them are deeply rooted social problems, and no mobile app can solve these issues immediately [65]. Although our app can provide some information about the underlying reasons in contexts such as this, it cannot go beyond to fix those problems.

Limitations and Future Work

Our paper does not make any causal claims about the social determinants of mental health (eg, we do not make claims that marriage causes mental frustration to young women). We only describe the findings of our interviews, and the factors we reported may not be causally related to mental health. Other hidden, unmeasured variables may play a role. However, we believe our reported factors will pave the way for future work that will dive deeply into investigating the underlying mechanisms of mental health, particularly the ones that came up for the first time in this study. We believe that follow-up work can then inform the theories of the mechanism of mental health. These theories can then be verified with actual field experiments to quantitatively find causal relationships.

Although we included students from 7 universities that offer different standards of education, future work should interact with students from even more diverse backgrounds. Our participants predominantly lived in 2 major cities, and students from universities in semiurban or rural areas might face other factors that have not been identified in this study. It would also be interesting to observe how this different demographic group struggles with the factors reported in this study and whether their perceptions about using a mobile app to monitor mental health would be any different.

Finally, there is a growing tendency among students to use social media platforms, such as Facebook Messenger, WhatsApp, or Emo, for having longer conversations. As the internet has started reaching even the most remote parts of the country [71], data plans and broadband internet are not as costly as before, and calling over the internet is a much cheaper option in many cases. Our participants also mentioned that they used

several platforms to contact friends or family members. Future studies can be designed on these internet platforms to monitor the mental health of the young, college-going population of Bangladesh in a more comprehensive way.

Conclusions

In this work, we aimed to progress the research toward identifying the social determinants of mental health among undergraduate students in Bangladesh. We designed and deployed an Android app among our participants to help them record and later reflect on the mobile phone conversations with their friends, family members, and academic or professional correspondences. Our app assisted the participants in pinpointing the exact relationship or factor that had a detrimental effect on their mental health. Although some of these factors have been

reported in previous studies conducted in the Western context, we identified several new factors, including religion and extended family affairs, which are pertinent to the society of Bangladesh. Although our app does not provide any recommendations from its side, some participants took independent measures to improve their mental health. However, in some cases (eg, extended family problems), despite identifying the problem, participants could not find an appropriate solution; however, they could better prepare themselves to cope with that specific issue. Taken together, our study provides useful perspectives on and insights into the mental health of undergraduate students in Bangladesh, and we hope our findings can help researchers design better solutions to improve the mental health of the younger population from this part of the world.

Acknowledgments

The project was supported by a Discovery Grant, Natural Sciences and Engineering Research Council, under grant number RGPIN-2018-0. The authors would like to thank Abdul Kawser Tushar (University of Toronto), Eftekher Ahmed (Chittagong University of Engineering and Technology), Mainul Islam (Verizon Media), and Md Sabir Hossain (Chittagong University of Engineering and Technology) for their contributions to the development of the app.

Conflicts of Interest

None declared.

References

1. Rotich O, Tugumisirize J. Mental health in developing countries: emphasis on Africa. Theseus. 2017. URL: <https://www.theseus.fi/handle/10024/133468> [accessed 2021-06-01]
2. Veen S, van Moorten R, van Durani W. Marriage Decision-Making: A family affair-Insight into marriage decision-making processes and social norms around child marriage in Larkana and Shikarpur in Sindh and Lodhran and Muzaffargarh in Punjab, Pakistan. Oxfam Novib. 2018. URL: <https://oxfamlibrary.openrepository.com/handle/10546/620535> [accessed 2021-05-01]
3. Nuri NN, Sarker M, Ahmed HU, Hossain MD, Beiersmann C, Jahn A. Pathways to care of patients with mental health problems in Bangladesh. *Int J Ment Health Syst* 2018;12:39 [FREE Full text] [doi: [10.1186/s13033-018-0218-y](https://doi.org/10.1186/s13033-018-0218-y)] [Medline: [30034515](https://pubmed.ncbi.nlm.nih.gov/30034515/)]
4. Hossain MD, Ahmed HU, Chowdhury WA, Niessen LW, Alam DS. Mental disorders in Bangladesh: A systematic review. *BMC Psychiatry* 2014 Jul 30;14:216 [FREE Full text] [doi: [10.1186/s12888-014-0216-9](https://doi.org/10.1186/s12888-014-0216-9)] [Medline: [25073970](https://pubmed.ncbi.nlm.nih.gov/25073970/)]
5. Firoz A, Karim M, Alam M, Rahman A, Zaman M. Prevalence, medical care, awareness and attitude towards mental illness in Bangladesh. *Bangladesh J Psychiatry* 2006;20(1):9-36 [FREE Full text]
6. Iqbal Y, Jahan R, Matin MR. Descriptive characteristics of callers to an emotional support and suicide prevention helpline in Bangladesh (first five years). *Asian J Psychiatr* 2019 Oct;45:63-65. [doi: [10.1016/j.ajp.2019.09.005](https://doi.org/10.1016/j.ajp.2019.09.005)] [Medline: [31520881](https://pubmed.ncbi.nlm.nih.gov/31520881/)]
7. Aggarwal NK. Applying mobile technologies to mental health service delivery in South Asia. *Asian J Psychiatr* 2012 Sep;5(3):225-230. [doi: [10.1016/j.ajp.2011.12.009](https://doi.org/10.1016/j.ajp.2011.12.009)] [Medline: [22981050](https://pubmed.ncbi.nlm.nih.gov/22981050/)]
8. Whatnall MC, Patterson AJ, Brookman S, Convery P, Swan C, Pease S, et al. Lifestyle behaviors and related health risk factors in a sample of Australian university students. *J Am Coll Health* 2020 Oct;68(7):734-741. [doi: [10.1080/07448481.2019.1611580](https://doi.org/10.1080/07448481.2019.1611580)] [Medline: [31140957](https://pubmed.ncbi.nlm.nih.gov/31140957/)]
9. Pendse S, Karusala N, Siddarth D, Gonsalves P, Mehrotra S, Naslund J, et al. Mental health in the global south: challenges and opportunities in HCI for development. In: Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies. 2019 Presented at: COMPASS '19: ACM SIGCAS Conference on Computing and Sustainable Societies; July 3 - 5, 2019; Accra, Ghana p. 22-36 URL: <https://doi.org/10.1145/3314344.3332483> [doi: [10.1145/3314344.3332483](https://doi.org/10.1145/3314344.3332483)]
10. Compton M, Shim R. The social determinants of mental health. *Focus* 2015 Oct;13(4):419-425. [doi: [10.1176/appi.focus.20150017](https://doi.org/10.1176/appi.focus.20150017)]
11. Stallard P, Porter J, Grist R. A smartphone app (Blueice) for young people who self-harm: Open phase 1 pre-post trial. *JMIR Mhealth Uhealth* 2018 Jan 30;6(1):e32 [FREE Full text] [doi: [10.2196/mhealth.8917](https://doi.org/10.2196/mhealth.8917)] [Medline: [29382632](https://pubmed.ncbi.nlm.nih.gov/29382632/)]
12. Shah MM, Ahmed S, Arafat SM. Demography and risk factors of suicide in Bangladesh: A six-month paper content analysis. *Psychiatry J* 2017;2017:3047025 [FREE Full text] [doi: [10.1155/2017/3047025](https://doi.org/10.1155/2017/3047025)] [Medline: [29130035](https://pubmed.ncbi.nlm.nih.gov/29130035/)]

13. Sultana N. Stress and depression among undergraduate medical students of Bangladesh. *Bangla J Med Educ* 2014 Mar 01;2(1):6-9 [[FREE Full text](#)] [doi: [10.3329/bjme.v2i1.18130](https://doi.org/10.3329/bjme.v2i1.18130)]
14. Mohiuddin AK. Drug addiction in Bangladesh: "A Consequence of Social Demoralization Rather Than Individual Flaws". *Int J Addict Res Ther* 2019;2(10):1005. [doi: [10.28933/ijart-2019-11-1005](https://doi.org/10.28933/ijart-2019-11-1005)]
15. Fair CC, Hamza A, Heller R. Who supports suicide terrorism in Bangladesh? What the data say. *SSRN Journal* 2020:2960630. [doi: [10.2139/ssrn.2960630](https://doi.org/10.2139/ssrn.2960630)]
16. Acharya JP, Acharya I, Waghrey D. A study on some of the common health effects of cell-phones amongst college students. *J Community Med Health Edu* 2013;3(4):1000214. [doi: [10.4172/2161-0711.1000214](https://doi.org/10.4172/2161-0711.1000214)]
17. Ahmed I, Perji KA. Mobile phone to youngsters: Necessity or addiction. *Inform Manag Busi Rev* 2011 May 15;2(5):229-238. [doi: [10.22610/imbr.v2i5.902](https://doi.org/10.22610/imbr.v2i5.902)]
18. Bredström A. Culture and context in mental health diagnosing: Scrutinizing the DSM-5 revision. *J Med Humanit* 2019 Sep;40(3):347-363 [[FREE Full text](#)] [doi: [10.1007/s10912-017-9501-1](https://doi.org/10.1007/s10912-017-9501-1)] [Medline: [29282590](https://pubmed.ncbi.nlm.nih.gov/29282590/)]
19. Jackson JS, Torres M, Caldwell CH, Neighbors HW, Nesse RM, Taylor RJ, et al. The National Survey of American Life: a study of racial, ethnic and cultural influences on mental disorders and mental health. *Int J Methods Psychiatr Res* 2004;13(4):196-207 [[FREE Full text](#)] [doi: [10.1002/mpr.177](https://doi.org/10.1002/mpr.177)] [Medline: [15719528](https://pubmed.ncbi.nlm.nih.gov/15719528/)]
20. Martin M. *The Social Epidemiology of Mental Disorders*—by E. Gartly Jaco, New York, Russell Sage Foundation, 1960, 228 pages, \$3.50. *Psychiatr Serv* 1961 Aug;12(8):40. [doi: [10.1176/ps.12.8.40](https://doi.org/10.1176/ps.12.8.40)]
21. Lewis-Fernández R, Kirmayer LJ. Cultural concepts of distress and psychiatric disorders: Understanding symptom experience and expression in context. *Transcult Psychiatry* 2019 Aug;56(4):786-803. [doi: [10.1177/1363461519861795](https://doi.org/10.1177/1363461519861795)] [Medline: [31347476](https://pubmed.ncbi.nlm.nih.gov/31347476/)]
22. Healthy people 2000 review. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics. 1993. URL: <https://www.cdc.gov/nchs/data/hp2000/hp2k01.pdf> [accessed 2021-01-06]
23. Bruffaerts R, Mortier P, Kiekens G, Auerbach RP, Cuijpers P, Demyttenaere K, et al. Mental health problems in college freshmen: Prevalence and academic functioning. *J Affect Disord* 2018 Jan 01;225:97-103. [doi: [10.1016/j.jad.2017.07.044](https://doi.org/10.1016/j.jad.2017.07.044)] [Medline: [28802728](https://pubmed.ncbi.nlm.nih.gov/28802728/)]
24. De Luca SM, Franklin C, Yueqi Y, Johnson S, Brownson C. The relationship between suicide ideation, behavioral health, and college academic performance. *Community Ment Health J* 2016 Jul;52(5):534-540. [doi: [10.1007/s10597-016-9987-4](https://doi.org/10.1007/s10597-016-9987-4)] [Medline: [26831304](https://pubmed.ncbi.nlm.nih.gov/26831304/)]
25. Kiekens G, Claes L, Demyttenaere K, Auerbach RP, Green JG, Kessler RC, et al. Lifetime and 12-month nonsuicidal self-injury and academic performance in college freshmen. *Suicide Life Threat Behav* 2016 Oct;46(5):563-576. [doi: [10.1111/sltb.12237](https://doi.org/10.1111/sltb.12237)] [Medline: [26954061](https://pubmed.ncbi.nlm.nih.gov/26954061/)]
26. Atri A, Sharma M, Cottrell R. Role of social support, hardiness, and acculturation as predictors of mental health among international students of Asian Indian origin. *Int Q Community Health Edu* 2006;27(1):59-73. [doi: [10.2190/IQ.27.1.e](https://doi.org/10.2190/IQ.27.1.e)] [Medline: [18039629](https://pubmed.ncbi.nlm.nih.gov/18039629/)]
27. Basu S, Garg S, Singh MM, Kohli C. Addiction-like behavior associated with mobile phone usage among medical students in Delhi. *Indian J Psychol Med* 2018;40(5):446-451 [[FREE Full text](#)] [doi: [10.4103/IJPSYM.IJPSYM_59_18](https://doi.org/10.4103/IJPSYM.IJPSYM_59_18)] [Medline: [30275620](https://pubmed.ncbi.nlm.nih.gov/30275620/)]
28. Mayer F, Santos I, Silveira PS, Lopes MH, de Souza AR, Campos EP, et al. Factors associated to depression and anxiety in medical students: A multicenter study. *BMC Med Educ* 2016 Oct 26;16(1):282 [[FREE Full text](#)] [doi: [10.1186/s12909-016-0791-1](https://doi.org/10.1186/s12909-016-0791-1)] [Medline: [27784316](https://pubmed.ncbi.nlm.nih.gov/27784316/)]
29. Silva RG, Figueiredo-Braga M. Evaluation of the relationships among happiness, stress, anxiety, and depression in pharmacy students. *Curr Pharm Teach Learn* 2018 Jul;10(7):903-910. [doi: [10.1016/j.cptl.2018.04.002](https://doi.org/10.1016/j.cptl.2018.04.002)] [Medline: [30236427](https://pubmed.ncbi.nlm.nih.gov/30236427/)]
30. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: Randomized controlled trial. *J Med Internet Res* 2012;14(3):e67 [[FREE Full text](#)] [doi: [10.2196/jmir.1858](https://doi.org/10.2196/jmir.1858)] [Medline: [22732135](https://pubmed.ncbi.nlm.nih.gov/22732135/)]
31. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. *Lancet* 2012 Apr;379(9826):1641-1652. [doi: [10.1016/s0140-6736\(12\)60149-4](https://doi.org/10.1016/s0140-6736(12)60149-4)]
32. Laidlaw A, McLellan J, Ozakinci G. Understanding undergraduate student perceptions of mental health, mental well-being and help-seeking behaviour. *Stud High Edu* 2015 Mar 25;41(12):2156-2168 [[FREE Full text](#)] [doi: [10.1080/03075079.2015.1026890](https://doi.org/10.1080/03075079.2015.1026890)]
33. Babu N, Hossain Z, Morales J, Vij S. Grandparents in Bangladesh, India, and Pakistan. In: *Grandparents in Cultural Context*. Milton, UK: Routledge; 2017:159-186.
34. Devine J, Hinks T, Naveed A. Happiness in Bangladesh: The role of religion and connectedness. *J Happiness Stud* 2017 Dec 16;20(2):351-371 [[FREE Full text](#)] [doi: [10.1007/s10902-017-9939-x](https://doi.org/10.1007/s10902-017-9939-x)]
35. Gardner K, Kotalova J. Belonging to others: Cultural construction of womanhood among muslims in a village in Bangladesh. *J Royal Anthropol Instit* 1995 Jun;1(2):445. [doi: [10.2307/3034740](https://doi.org/10.2307/3034740)]
36. Gankovsky Y. The social structure of society in the People's Republic of Bangladesh. *Asian Survey* 1974 Mar;14(3):220-230 [[FREE Full text](#)] [doi: [10.1525/as.1974.14.3.01p04282](https://doi.org/10.1525/as.1974.14.3.01p04282)]

37. Gulliford L, Morgan B, Hemming E, Abbott J. Gratitude, self-monitoring and social intelligence: A prosocial relationship? *Curr Psychol* 2019 Jul 19;38(4):1021-1032 [FREE Full text] [doi: [10.1007/s12144-019-00330-w](https://doi.org/10.1007/s12144-019-00330-w)]
38. Fudolig MI, Bhattacharya K, Monsivais D, Jo H, Kaski K. Link-centric analysis of variation by demographics in mobile phone communication patterns. *PLoS One* 2020;15(1):e0227037 [FREE Full text] [doi: [10.1371/journal.pone.0227037](https://doi.org/10.1371/journal.pone.0227037)] [Medline: [31899785](https://pubmed.ncbi.nlm.nih.gov/31899785/)]
39. Bangladesh Telecommunication Regulatory Commission (BTRC). URL: <http://www.btrc.gov.bd/content/mobile-phone-subscribers-bangladesh-march-2019/> [accessed 2020-12-17]
40. Population and Housing Census 2011. Bangladesh Bureau of Statistics. 2011. URL: <http://www.bbs.gov.bd/site/page/47856ad0-7e1c-4aab-bd78-892733bc06eb/Population-and-Housing-Census> [accessed 2020-12-17]
41. Agarwal N. Being a witness: Using kinesthetic empathy as a reflective tool for clients. Expressive Therapies Capstone Theses. 2019. URL: https://digitalcommons.lesley.edu/expressive_theses/195 [accessed 2021-01-05]
42. Collins M, Long R, Page A, Popay J, Lobban F. Using the public involvement impact assessment framework to assess the impact of public involvement in a mental health research context: A reflective case study. *Health Expect* 2018 Dec;21(6):950-963 [FREE Full text] [doi: [10.1111/hex.12688](https://doi.org/10.1111/hex.12688)] [Medline: [29696740](https://pubmed.ncbi.nlm.nih.gov/29696740/)]
43. Wetzel M, Maloch B, Hoffman J. Retrospective video analysis: A reflective tool for teachers and teacher educators. *Read Teach* 2016 Nov 18;70(5):533-542 [FREE Full text] [doi: [10.1002/trtr.1550](https://doi.org/10.1002/trtr.1550)]
44. Desmet P, Romero N, Vastenburger MH. Mood measurement with Pick-A-Mood: Review of current methods and design of a pictorial self-report scale. *J Design Res* 2016;14(3):241-279 [FREE Full text] [doi: [10.1504/jdr.2016.10000563](https://doi.org/10.1504/jdr.2016.10000563)]
45. Kauer SD, Reid SC, Sanci L, Patton GC. Investigating the utility of mobile phones for collecting data about adolescent alcohol use and related mood, stress and coping behaviours: Lessons and recommendations. *Drug Alcohol Rev* 2009 Jan;28(1):25-30. [doi: [10.1111/j.1465-3362.2008.00002.x](https://doi.org/10.1111/j.1465-3362.2008.00002.x)] [Medline: [19320672](https://pubmed.ncbi.nlm.nih.gov/19320672/)]
46. Pollak J, Adams P, Gay G. PAM: a photographic affect meter for frequent, in situ measurement of affect. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2011; Vancouver BC Canada p. 725-734 URL: <https://doi.org/10.1145/1978942.1979047> [doi: [10.1145/1978942.1979047](https://doi.org/10.1145/1978942.1979047)]
47. 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/)]
48. 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/)]
49. Caldeira C, Chen Y, Chan L, Pham V, Chen Y, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. *AMIA Annu Symp Proc* 2017;2017:495-504 [FREE Full text] [Medline: [29854114](https://pubmed.ncbi.nlm.nih.gov/29854114/)]
50. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE J Biomed Health Inform* 2015 Jan;19(1):140-148. [doi: [10.1109/JBHI.2014.2343154](https://doi.org/10.1109/JBHI.2014.2343154)] [Medline: [25073181](https://pubmed.ncbi.nlm.nih.gov/25073181/)]
51. Huisman G, Van Hout M. The development of a graphical emotion measurement instrument using caricatured expressions: The LEMtool. In: Proceedings of the Emotion in HCI – Designing for People Conference. 2010 Presented at: Emotion in HCI – Designing for People Conference; 2008; Rostock, Germany p. 5-8 URL: https://www.researchgate.net/publication/259937797_The_development_of_a_graphical_emotion_measurement_instrument_using_caricatured_expressions_the_LEMtool
52. Sánchez J, Hernández N, Penagos J, Ostróvska Y. Conveying mood and emotion in instant messaging by using a two-dimensional model for affective states. In: Proceedings of VII Brazilian symposium on Human factors in computing systems. 2006 Presented at: IHC'06: VII Symposium on Human Factors in Computer Systems; November 19 - 22, 2006; Natal, RN Brazil p. 66-72 URL: <https://doi.org/10.1145/1298023.1298033> [doi: [10.1145/1298023.1298033](https://doi.org/10.1145/1298023.1298033)]
53. Goodman LA. Snowball sampling. *Ann Math Statist* 1961 Mar;32(1):148-170. [doi: [10.1214/aoms/1177705148](https://doi.org/10.1214/aoms/1177705148)]
54. Islam M, Alam M. Webometric study of private universities in Bangladesh. *Malay J Library Inform Sci* 2017;16(2):115-126.
55. Laws of Bangladesh. URL: <http://bdlaws.minlaw.gov.bd/> [accessed 2020-12-27]
56. Russell JA. A circumplex model of affect. *J Personal Soc Psychol* 1980;39(6):1161-1178. [doi: [10.1037/h0077714](https://doi.org/10.1037/h0077714)]
57. Hetrick SE, Robinson J, Burge E, Blandon R, Mobilio B, Rice SM, et al. Youth codesign of a mobile phone app to facilitate self-monitoring and management of mood symptoms in young people with major depression, suicidal ideation, and self-harm. *JMIR Ment Health* 2018 Jan 23;5(1):e9 [FREE Full text] [doi: [10.2196/mental.9041](https://doi.org/10.2196/mental.9041)] [Medline: [29362208](https://pubmed.ncbi.nlm.nih.gov/29362208/)]
58. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. California: SAGE Publications; Apr 16, 1998.
59. Ahmed AK, Ferdous L, Sharif AM. Career counseling at the universities: The Bangladesh scenario. *Euro Scient J* 2017 Oct 31;13(28):381. [doi: [10.19044/esj.2017.v13n28p381](https://doi.org/10.19044/esj.2017.v13n28p381)]
60. Hossain MI. Career choice decisions by business students: A study on some private universities in bangladesh. *Dspace Repository*. 2019. URL: <http://dspace.daffodilvarsity.edu.bd:8080/handle/123456789/3938> [accessed 2021-05-01]
61. Thorn P, Raj J. A culture of coaching: achieving peak performance of individuals and teams in academic health centers. *Academic Medicine*. 2012. URL: <http://hdl.handle.net/10504/61214> [accessed 2021-06-01]
62. Eisenberg D, Golberstein E, Hunt J. Mental health and academic success in college. *BE J Econ Anal Policy* 2009 Sep 15;9(1):1-37 [FREE Full text] [doi: [10.2202/1935-1682.2191](https://doi.org/10.2202/1935-1682.2191)]

63. Thomas S, Heinrich S, Kühnlein A, Radon K. The association between socioeconomic status and exposure to mobile telecommunication networks in children and adolescents. *Bioelectromagnetics* 2010 Jan;31(1):20-27. [doi: [10.1002/bem.20522](https://doi.org/10.1002/bem.20522)] [Medline: [19598181](https://pubmed.ncbi.nlm.nih.gov/19598181/)]
64. Islam M, Alim M, Hawlader S, Mohiuddin M, Sarker A. Assessment and analysis of the overall situation of women and children: Bangladesh scenario. *Int J Public Admin Policy Res* 2019;4(1):26-30 [[FREE Full text](#)]
65. Sultana S, Guimbretière F, Sengers P, Dell N. Design within a patriarchal society: Opportunities and challenges in designing for rural women in Bangladesh. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; April 21 - 26, 2018; Montreal QC Canada p. 1-13. [doi: [10.1145/3173574.3174110](https://doi.org/10.1145/3173574.3174110)]
66. Lee-Rife S, Malhotra A, Warner A, Glinski A. What works to prevent child marriage: A review of the evidence. *Stud Fam Plann* 2012 Dec;43(4):287-303. [doi: [10.1111/j.1728-4465.2012.00327.x](https://doi.org/10.1111/j.1728-4465.2012.00327.x)] [Medline: [23239248](https://pubmed.ncbi.nlm.nih.gov/23239248/)]
67. The state of the world's children 2016: A fair chance for every child. UNICEF. 2016. URL: <https://www.unicef.org/publications/files/unicef????????????2016????????????> [accessed 2020-12-19]
68. Santhya KG, Ram U, Acharya R, Jejeebhoy SJ, Ram F, Singh A. Associations between early marriage and young women's marital and reproductive health outcomes: Evidence from India. *Int Perspect Sex Reprod Health* 2010 Sep;36(3):132-139. [doi: [10.1363/ipsrh.36.132.10](https://doi.org/10.1363/ipsrh.36.132.10)] [Medline: [20880798](https://pubmed.ncbi.nlm.nih.gov/20880798/)]
69. The state religion. URL: <https://bdlaws.minlaw.gov.bd> [accessed 2020-12-19]
70. Memon M, Bhatti S, Mahoto N. A digital diary: Remembering the past using the present context. *Mehran Uni Res J Eng Technol* 2016;35(2):275-286 [[FREE Full text](#)] [doi: [10.3316/informit.133237135502851](https://doi.org/10.3316/informit.133237135502851)]
71. Hassan T, Alam MM, Wahab A, Hawlader MD. Prevalence and associated factors of internet addiction among young adults in Bangladesh. *J Egypt Public Health Assoc* 2020 Jan 29;95(1):3 [[FREE Full text](#)] [doi: [10.1186/s42506-019-0032-7](https://doi.org/10.1186/s42506-019-0032-7)] [Medline: [32813097](https://pubmed.ncbi.nlm.nih.gov/32813097/)]

Abbreviations

BUET: Bangladesh University of Engineering and Technology

HWC: How Was the Call?

Edited by G Eysenbach; submitted 12.01.21; peer-reviewed by A Das, A Andy, J Smyth; comments to author 18.02.21; revised version received 10.05.21; accepted 01.08.21; published 02.11.21.

Please cite as:

Bhattacharjee A, Haque SMT, Hady MA, Alam SMR, Rabbi M, Kabir MA, Ahmed SI

Understanding the Social Determinants of Mental Health of Undergraduate Students in Bangladesh: Interview Study

JMIR Form Res 2021;5(11):e27114

URL: <https://formative.jmir.org/2021/11/e27114>

doi: [10.2196/27114](https://doi.org/10.2196/27114)

PMID: [34726609](https://pubmed.ncbi.nlm.nih.gov/34726609/)

©Ananya Bhattacharjee, S M Taiabul Haque, Md Abdul Hady, S M Raihanul Alam, Mashfiqui Rabbi, Muhammad Ashad Kabir, Syed Ishtiaque Ahmed. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 02.11.2021. 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 App to Enhance Behavioral Activation Treatment for Substance Use Disorder: App Design, Use, and Integration Into Treatment in the Context of a Randomized Controlled Trial

Catherine E Paquette¹, MPS, MA; Dillon T Rubalcava¹, BSc; Yun Chen¹, PhD; Deepika Anand^{1,2}, PhD; Stacey B Daughters¹, PhD

¹Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

²CBT Center of Chicago, Chicago, IL, United States

Corresponding Author:

Stacey B Daughters, PhD

Department of Psychology and Neuroscience

University of North Carolina at Chapel Hill

235 E. Cameron Avenue

Chapel Hill, NC, 27599-3270

United States

Phone: 1 919 962 9924

Email: daughter@unc.edu

Abstract

Background: Group-based formats typically used in low-resource substance use disorder (SUD) treatment settings result in little individual attention to help reinforce and guide skill use, which may contribute to poor posttreatment outcomes. Smartphone apps offer a convenient, user-friendly, and cost-effective tool that can extend the reach of effective SUD treatments. A smartphone app was developed and integrated into a group-based, brief behavioral activation (BA) treatment for SUD to increase engagement in treatment skills outside clinician-administered sessions.

Objective: This study aims to describe the features of the app and its use and integration into treatment, report the participants' self-reported feasibility and acceptability of the app, and discuss challenges and provide recommendations for future smartphone app integration into behavioral treatments for SUD.

Methods: A total of 56 individuals recruited from intensive outpatient SUD treatment received a smartphone-enhanced BA treatment, the Life Enhancement Treatment for Substance Use. Self-reported weekly app use and reasons for nonuse were assessed at posttreatment and at 1- and 3-month follow-ups. In addition, 2-tailed *t* tests and chi-square tests compared the self-reported use of each app component and overall app use over time.

Results: Participant feedback suggested that the integration of the smartphone app into the Life Enhancement Treatment for Substance Use was feasible and well accepted, and participants found the app useful for planning value-based activities outside of sessions. Self-reported app engagement decreased over the follow-up period: 72% (39/54) of participants reported using the app at posttreatment, decreasing to 69% (37/54) at the 1-month follow-up and 37% (20/54) at the 3-month follow-up. Participants reported forgetting to use the app as a primary reason for nonuse.

Conclusions: This study provides support for the feasibility and acceptability of smartphone-enhanced BA treatment, offering promise for future research testing the integration of technology into SUD treatment. Design decisions may help streamline smartphone integration into treatment, for example, allowing participants to download the treatment app on their own phones or use a low-cost study smartphone (or offering both options). Long-term app engagement may be increased via built-in reminders, alerts, and in-app messages.

Trial Registration: ClinicalTrials.gov NCT02707887; <https://clinicaltrials.gov/ct2/show/study/NCT02707887>

(*JMIR Form Res* 2021;5(11):e25749) doi:[10.2196/25749](https://doi.org/10.2196/25749)

KEYWORDS

substance use disorder; smartphone app; mHealth; behavioral activation; mobile phone

Introduction

Background

Limited access to evidence-based substance use disorder (SUD) treatment is a pervasive problem in the United States. Of the 20.3 million Americans who experienced an SUD in the past year, only approximately 12% received treatment at a specialty facility [1]. Treatment providers at these facilities often lack extensive training in evidence-based interventions [2]. Most SUD treatment is provided in a group-based format [3], in which participants receive little individual attention, and group-based peer support models such as 12-Step Facilitation are the primary method of care at many treatment centers [2,4]. Ongoing monitoring and aftercare for SUD treatment is rare [5,6], although prolonged treatment engagement is associated with better outcomes [7,8]. There is a clear need for evidence-based treatments that prioritize ongoing engagement and are accessible within low-resource SUD treatment settings (including treatment being provided by practitioners without graduate-level training).

Behavioral Activation for Substance Use

The Life Enhancement Treatment for Substance Use (LETS ACT) [9] was developed to address this need. LETS ACT is a behavioral activation (BA) treatment, which aims to increase substance-free environmental reinforcement through the planning and execution of value-based activities. SUDs are characterized by reward deficits and a loss of drug-free positive reinforcement, resulting in decreased engagement in naturally rewarding (ie, drug-free) activities [10]. BA targets this lack of environmental reward by helping participants increase their daily engagement in activities that are enjoyable and important and thus provides opportunities for positive reinforcement. It has shown efficacy when delivered by individuals without professional training in psychotherapy [11], making it accessible for implementation in low-resource SUD treatment settings. LETS ACT, provided as a supplement to inpatient SUD treatment, has demonstrated effectiveness in reducing depressive symptoms [12] and rates of treatment dropout [13] as well as increasing rates of abstinence and decreasing substance-related consequences up to 1 year posttreatment [14].

Despite positive initial findings, there continues to be room to improve posttreatment outcomes by increasing out-of-session treatment engagement. Although significantly lower than a contact-matched control condition, more than 50% of LETS ACT participants reported using substances by 3 months posttreatment [14]. There is little individual attention to help reinforce and guide BA skills outside of the 6 group sessions. Research suggests that homework compliance is predictive of better treatment outcomes across psychotherapy modalities [15], including in cognitive behavioral treatments such as BA [16,17]. Thus, bolstering out-of-session engagement could be one way to improve posttreatment outcomes.

Smartphone-Enhanced Treatment

Integrating smartphones into therapy is a promising strategy for increasing engagement outside group-based BA sessions. Features such as built-in guidance, prompts, and reminders can assist individuals in completing homework in a manner

compliant with treatment guidelines, for example, by reminding participants to link planned activities to specific values. Users of smartphone apps for addiction recovery frequently cite the portability of apps as an advantage as well as their discreet nature [18]. Participants in BA treatments for depression have noted that smartphone-based treatments are more accessible to their everyday lives [19], which is key given the importance of daily activity planning in BA. Smartphone apps allow for quick access to skills learned in therapy within naturalistic settings, such as reviewing a plan for healthy coping behaviors in contexts with a high risk of substance use. Smartphone apps can also provide a cost-effective way of engaging patients [20], which makes them promising for low-resource, group-based SUD treatment settings.

Current research suggests that interventions involving smartphone apps are feasible and well accepted among individuals with SUDs. Research in SUD treatment samples demonstrates high rates of smartphone ownership and use, similar to the general population [21-23], and recent studies have found high overall acceptability of mobile health interventions for SUD [24]. Previous smartphone-based interventions have demonstrated positive effects on substance-related behavioral changes [24,25]. However, very few stand-alone apps use evidence-based interventions such as cognitive behavioral therapy [26], and research testing the integration of smartphone technology into established SUD treatments is lacking. Indeed, in a recent systematic review examining smartphone-based treatments for psychiatric diagnoses, only 2 of 27 studies identified by the review examined SUD interventions [27]. Of these, only 1 assessed a smartphone-enhanced treatment as opposed to a stand-alone app-based intervention, finding that the smartphone-enhanced treatment was associated with greater reductions in substance use [28]. Thus, there is a clear need for research examining the feasibility and utility of integrating smartphones into evidence-based SUD treatment.

Study Objectives

This study reports feasibility data from a trial (NCT02707887) testing the effectiveness of a smartphone-enhanced BA treatment for SUD (smartphone-enhanced LETS ACT). The aims of the study are to (1) describe the features of the app and its use and integration into treatment, (2) report participants' self-reported feasibility and acceptability of the app, and (3) discuss challenges and provide recommendations for future smartphone app integration into behavioral treatments for SUD.

Methods

Development of the LETS ACT App

The LETS ACT app was designed to largely reflect the paper treatment materials used in previous studies of LETS ACT, with a number of added features to facilitate theory-driven treatment engagement. In the development phase, the research team drew from prior research and consultation with researchers and clinicians with expertise in SUD treatment and the development of technology to enhance behavioral interventions. The design included app features intended to address some of the limitations of paper materials (eg, providing in-app suggestions for

improving homework compliance based on the user's weekly progress). Furthermore, the app was designed to collect daily mood and substance use data. The final app was developed through an iterative piloting process, which included testing a web-based version before piloting the app with individuals in inpatient SUD treatment.

Design of This Study

The data presented here come from a single-site, 3-arm trial conducted at an intensive outpatient SUD treatment center in Raleigh, North Carolina, comparing smartphone-enhanced BA with standard BA and treatment as usual (TAU). The focus of this analysis is to determine the feasibility and acceptability of the smartphone-enhanced treatment condition before a future report of the main outcomes of the parent trial. All participants received TAU. A total of 65 participants were randomized to smartphone-enhanced LETS ACT and attended at least one session of treatment; of these, 56 attended a second session and received a smartphone. Data for this study were collected at the pretreatment assessment, posttreatment, and at 1- and 3-month posttreatment follow-ups (FU1 and FU3). All study procedures were approved by the institutional review board.

Sample and Recruitment

Patients at the outpatient facility were primarily low-income individuals with a range of SUD diagnoses who voluntarily enrolled in the treatment. Patients were recruited by the research team weekly through announcements at the end of the TAU treatment groups and by approaching individuals after these groups were released. Interested individuals were assessed for eligibility, provided informed consent, and completed the pretreatment assessment. Randomization occurred at the group level using a computerized urn randomization program, and participants were blinded to the condition (ie, participants were recruited in waves and were unaware of the 3 arms of the trial). Study exclusion criteria were (1) age >65 or <18 years, (2) less than fifth grade English reading level (ie, score <42 on the Wide Range Achievement Test), (3) current impairment due to psychotic symptoms, (4) completion of >6 weeks of TAU, and (5) inability to give written informed consent to participate. Following treatment, participants completed FU assessments

at the outpatient treatment facility or a public location with adequate privacy (eg, public library).

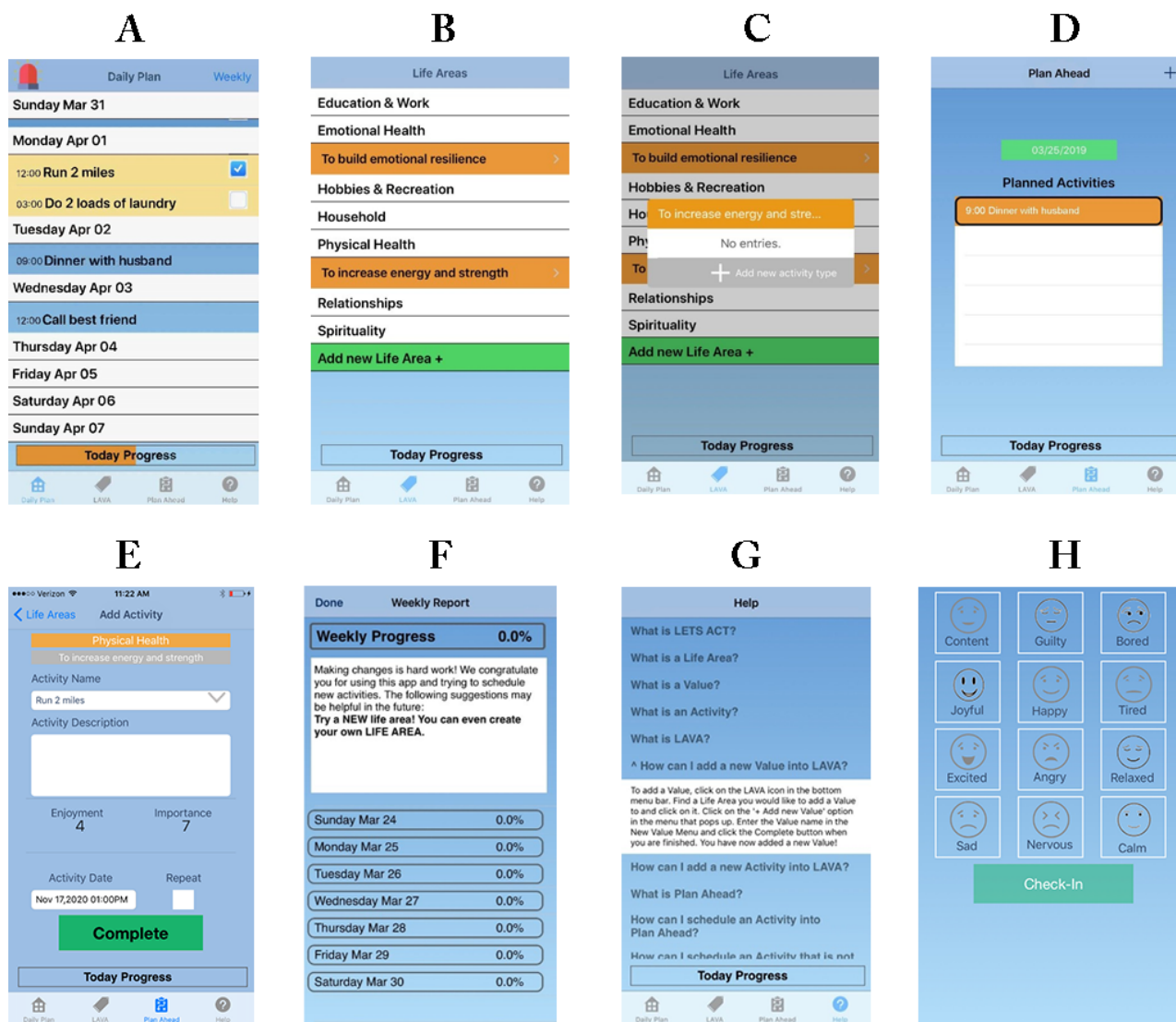
Intervention

Smartphone-Enhanced LETS ACT

The smartphone app was developed as an adjunct to LETS ACT [9]. LETS-ACT-SE is provided in small groups of 6 or fewer participants twice weekly over 3 weeks (6 sessions total). Each session begins with a discussion of the treatment rationale, including describing the cycle of negative mood, urges, and maladaptive behaviors (eg, substance use) and eliciting participant examples of how this cycle is experienced. Participants learn that the goal of treatment is to break this cycle by engaging in healthy, rewarding behaviors. They are taught that when an individual regularly engages in activities that generate a sense of enjoyment or accomplishment (or both), they are less likely to have urges to use substances or engage in other maladaptive behaviors in response to difficult emotions. Following the treatment rationale, participants record daily activities and rate them on enjoyment and importance to identify patterns of inactivation and opportunities to increase positive reinforcement.

Next, emphasis shifts to an activity called Life Areas, Values, and Activities (LAVA). LAVA involves identifying activities associated with specific values and life areas (eg, education and work, emotional health, hobbies and recreation, and relationships). Participants are guided through the LAVA activity by selecting a life area that is important to them (eg, physical health), then identifying a value they hold related to that life area by answering the question, "What is important to me within this life area?" (eg, "It is important to me to increase energy and strength"). Participants then generate specific, measurable activities aligned with their values, with an emphasis on balancing enjoyable and important activities (eg, "In order to have energy and strength [value], I will walk in the park for 30 minutes [activity]"). Earlier sessions focus on tracking daily activities and creating LAVA lists. Later sessions shift focus to planning and implementing these activities in a daily plan (Figure 1A), problem-solving challenges to adherence, and posttreatment planning.

Figure 1. Screenshots of the LETS ACT app. A: Daily Plan with completed activity; B: Life Areas, Values, and Activities; C: Activity prompt in Life Areas, Values, and Activities feature; D: Plan Ahead; E: Rating enjoyment and importance; F: Weekly Progress; G: Help page; H: Mood ratings.



Participants are given home practice assignments after each session, which include instructions for the continued use of each component. For example, after the introduction of LAVA in session 2, participants are asked to record at least one value and activity for their chosen life areas. After the introduction of the daily plan in session 3, participants are asked to plan and complete at least one activity per day for the remainder of the treatment. Participants are encouraged to continue planning and completing activities after the completion of treatment using their smartphones; however, they are not given any specific assignments to complete during the FU period.

Participants in LETS ACT-SE were provided with Apple iPhone 6 smartphones with the LETS ACT app predownloaded during the second treatment session. Phone plans were set up and paid for by the research study; plans included unlimited calls and text messages and 4 GB of wireless data per month. The intent of this service was for participants to use their phone for regular use, thus allowing the research team to assess the feasibility of the app on a personal use device. This ensured that all participants had consistent access to the LETS ACT app (which was programmed specifically for the iPhone to limit

development costs) as well as to reliable internet access throughout the study duration, allowing for ongoing data collection. At this time, they were given a brief introduction to the smartphones and LETS ACT app as well as a packet of information about basic features of the phones and instructions for use (eg, how to change settings). Participants absent in session 2 were given the phone and instructions at the next treatment session attended. Participants were introduced to each app component during the sessions, with a quick therapist-led tutorial followed by in-session practice. Participants were asked to use the smartphone app to record their homework. They were informed that the smartphones were theirs to use until their FU3 appointment, at which time they returned their phones to the research team and received monetary compensation. Individuals who lost their smartphone or forgot to bring it to treatment were provided with equivalent paper forms.

Treatment as Usual

All study participants were enrolled in a substance use disorder intensive outpatient program, in which treatment is based on the matrix model of intensive outpatient treatment [29]. The program included group therapy (average of 8-10 patients per

group) for 3 hours per day, 3 days per week for 12 weeks, as well as weekly individual appointments with a case manager and up to 2 optional individual counseling sessions per week. Although group sessions do not have a set curriculum, they typically include individual check-ins, psychoeducation (eg, related to relapse prevention), and time to verbally process and share. Urine drug testing is implemented throughout the treatment, and positive drug tests are openly discussed within group sessions. Continued use of substances (aside from nicotine) is grounds for dismissal from the program.

App Design and Components

Key components of the LETS ACT app include the LAVA library, Plan Ahead or Daily Plan, Weekly Progress, and Emergency button, accessible via icons on the home screen of the app (this is also the Plan Ahead screen; [Figure 1A](#)). Additional features include a Help page and a data collection mechanism for mood and substance use tracking.

Life Areas, Values, and Activities

The LAVA feature ([Figure 1B](#)) guides the user through the 3 steps of selecting value-based activities (described previously), reflecting the way the LAVA activity is taught during treatment. This was designed to increase the likelihood that the selected activities were value based. After selecting the LAVA icon, the user is presented with a list of life areas and an option to add a new life area. By tapping on a life area, the user is prompted to add a new value. Once complete, the value is listed in orange underneath the associated life area on the LAVA screen, and the participant can add activities by selecting the value ([Figure 1C](#)). Users can enter multiple values within each life area and perform multiple activities under each value.

Plan Ahead and Daily Plan

Planning value-based activities is central to the LETS ACT treatment, and the app includes 2 features that assist with this. The Plan Ahead feature ([Figure 1D](#)) allows the user to schedule value-based activities for specific days and times. By tapping a plus sign, the user is brought to the list of life areas, where they can either select an activity previously entered in the LAVA feature or enter a new activity (ie, by first selecting a life area, then entering a value and corresponding activity). Once an activity is selected, the user is prompted to rate the activity on enjoyment and importance ([Figure 1E](#)). Finally, the user can select a specific date and time to complete the activity, with the option of repeating the activity daily and/or weekly. The activity is entered into the user's daily plan ([Figure 1A](#)), and upcoming planned activities are listed by date on the Plan Ahead screen ([Figure 1D](#)).

The Daily Plan feature is the home screen of the app ([Figure 1A](#)). Activities planned for the coming week are listed by day, and users can mark activities complete by checking a box at or

after the assigned completion time (until midnight on the same day). Activities not marked complete by midnight are recorded as incomplete and removed from the Plan Ahead and Daily Plan screens.

Weekly Progress

On the home screen, an option in the top right corner allows the user to view their weekly progress, that is, the percentage of planned activities completed in the previous week ([Figure 1F](#)). An overall percentage is displayed at the top of the screen and the percentages for each day of the week underneath. In addition, this screen displays feedback and suggestions based on the user's progress. The Daily Plan screen also features a *Today Progress* bar that fills in with orange based on the percentage of completed activities for the current day and turns green when 100% of activities have been completed ([Figure 1A](#)).

Emergency Button for High-risk Situations

The Emergency button appears as a red siren at the top left of the Daily Plan screen ([Figure 1A](#)) and allows the user to create a list of *emergency activities* or healthy coping behaviors they can use while experiencing difficult emotions and/or urges to use substances. The emergency screen lists the user's emergency activities, which can be quickly added to the daily plan by selecting an activity title. Once selected, activities are marked complete at that specific date and time.

Help Icon

The Help icon brings the user to a page ([Figure 1G](#)) with a list of frequently asked questions and their answers, including information about the primary treatment components (eg, "What is a Value?") and instructions for using the app features (eg, "How can I schedule an activity into Plan Ahead?").

Research Functions

On opening the app for the first time each day, the user is prompted to rate their current mood ([Figure 1H](#)) and report any substance use (except for nicotine) since they last opened the app.

Measures and Outcome Variables

A questionnaire administered at pretreatment assessed smartphone ownership, use, and likelihood of using a smartphone for a research study. In addition, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5) diagnoses for mood, anxiety, and SUD were assessed at pretreatment using the Mini-International Neuropsychiatric Interview [30]. Sociodemographic information (including age and education level) was assessed at all time points. Detailed information about the measures and outcome variables related to self-reported app use and app component feedback is provided in [Table 1](#) and described below.

Table 1. Measures and outcome variables.

Construct	Variable description	Scale or possible value range	Time points
Past-week use of app components ^a	Multiple (all past week): 1. # days created new life areas, values, and activities using the LAVA ^b library 2. # days scheduled ≥ 1 activity into daily plan using Plan Ahead 3. # days used the Emergency Button ≥ 1 time 4. # days viewed Weekly Progress 5. # days viewed Help icon	0-7 days	PT ^c
Average weekly use of app components in past month ^a	Multiple (all past month): 1. Average # days/week entered life areas, values, and activities using LAVA icon 2. Average # days/week used Daily Plan icon 3. Average # activities planned >1 week in advance using Plan Ahead icon 4. Average # days/week used Emergency button 5. Average # days/week viewed Weekly Progress 6. Average # days/week viewed Help page	0-7 days	FU1 ^d and FU3 ^e
Any app use ^a	Whether participant reported using any (ie, one or more) app component at least 1 day per week since the previous assessment	Yes or no	PT, FU1, FU3
App component usefulness ^f	Degree to which participant agrees each component was a useful part of treatment: 1. LAVA library 2. Plan Ahead 3. Emergency 4. Weekly Progress 5. Help	Scale of 1 (strongly disagree) to 5 (strongly agree)	PT
Reasons for not using specific app components ^f	Reasons for not using: LAVA library Plan Ahead Emergency Weekly Progress Help	Can select all that apply from a list of reasons, select <i>other</i> , or indicate that it does not apply because the participant did use that component	PT
Reasons for low weekly app use ^f	Reasons for not using the app at least 3 times a week	Can select all that apply from a list of reasons, select <i>other</i> , or indicate that it does not apply because the participant did use the app at least three times per week	PT

^aDescribed under Self-reported Use of App Components section.

^bLAVA: Life Areas, Values, and Activities.

^cPT: posttreatment.

^dFU1: 1-month follow-up.

^eFU2: 3-month follow-up.

^fDescribed under App Component Usefulness and Reasons for Not Using section.

Self-reported Use of App Components

A questionnaire administered at posttreatment assessed participants' self-reported app engagement during the past week. Participants indicated the number of days in the past week that they used each treatment component outside of the treatment sessions. At FU assessments (ie, FU1 and FU3), participants were given a similar questionnaire that assessed engagement

with the app components during the past month. This included the average number of days per week that the participant used each component of the LETS ACT app and details about their use (eg, the number of activities scheduled and completed and the number of days per week with at least one scheduled activity; Table 1). A dichotomous variable representing any app use at each time point was calculated, such that participants were coded as having used the app at each time point if they reported using

one or more app components on 1 or more days per week since the previous assessment.

App Component Usefulness and Reasons for Not Using

A questionnaire administered at posttreatment assessed participant feedback about the treatment and its components (Table 1). For each component, participants rated the degree to which they agreed that the component was a useful part of the treatment on a scale of 1 (strongly disagree) to 5 (strongly agree). The measure also included questions assessing the reasons for *not* using each component. Participants could choose any applicable reasons from a list (eg, did not remember to use the feature, did not think it would be helpful, and difficult to understand how to use it), record their own reason under *other*, or indicate that the question did not apply to them because they did use that component. An additional question inquired as to any reasons participants did not use the app at least three times per week. Similarly, participants could select from a list of reasons, select *other*, or indicate that it did not apply because they did use the app at least three times per week. Finally, the questionnaire included an open-ended question eliciting feedback about the smartphone-enhanced treatment overall, including the degree to which it was useful and any suggestions for improvement.

Statistical Analyses

Data were analyzed using SPSS (version 25.0, IBM Corp). First, descriptive statistics were calculated for all variables used in subsequent statistical analyses. This included means, SDs, and ranges for continuous variables and percentages for all categorical variables. To examine the feasibility and acceptability of the LETS ACT app, summary statistics (eg, mean, median, and SD) were calculated to characterize participant ratings regarding the usefulness of each app component, as well as self-reported engagement with each component (ie, past-week use of app components at posttreatment and average weekly use of app components in the past month at FU1 and FU3). Chi-square tests were used to examine differences in the proportions of participants who reported any app use at each time point. Two-tailed paired-sample *t* tests were used to examine the differences in the mean ratings for app usefulness across components. The reasons for not using each app component and reasons for low weekly use of the app were also summarized.

Results

Sample Characteristics

Of the 56 participants who received a smartphone, 21 (38%) were women. Overall, 61% (34/56) of participants were White and 38% (21/56) were Black. The average age was 42.4 (SD 10.5; range 24–62) years, and the participants had an average of 12.1 (SD 3.0; range 1–21) years of education. In terms of substance use, participants reported an average of 3.8 (SD 6.9; range 0–30) days of substance use in the past 30 days at pretreatment. For DSM-5 SUDs, 73% (41/56) of participants met the criteria for alcohol use disorder, 59% (33/56) met the criteria for cocaine use disorder, 45% (25/56) met the criteria for opioid use disorder, and 30% (17/56) met the criteria for

cannabis use disorder. Regarding psychiatric comorbidity at pretreatment, 11% (6/56) of participants met the DSM-5 criteria for a current major depressive episode, and the average score on the Beck Depression Inventory at pretreatment was 12 (SD 11.3; range 0–51), reflecting minimal depressive symptoms. A total of 29% (16/56) of participants met the criteria for at least one DSM-5 anxiety disorder, including 18% (10/56) for social anxiety disorder, 11% (6/56) for agoraphobia and generalized anxiety disorder, 9% (5/56) for panic disorder, and 7% (4/56) for obsessive-compulsive disorder. Furthermore, 11% (6/56) met the criteria for bipolar I disorder, and 9% (5/56) met the criteria for posttraumatic stress disorder.

Smartphone Interruptions

Of the 56 participants who received a smartphone, 54 (96%) were retained in the study through the 3-month FU assessment. One participant withdrew from the study at posttreatment because of having a busy work schedule, and a second died (of non-study-related causes) between the posttreatment and 1-month FU assessments. Overall, 27% (15/56) of participants reported at least one interruption in their ability to use their smartphone during or after treatment up to the 3-month FU, including phone lost or stolen ($n=8$), inability to access phones due to incarceration ($n=3$), and other issues (eg, technical issues with the phone; $n=4$).

Smartphone Ownership and Use at Pretreatment

Among participants who received a smartphone, pretreatment data from those who reported their current smartphone ownership ($n=38$) indicated that 79% (30/38) owned a smartphone they could use daily. In total, 14% (5/37) reported that they owned an iPhone, whereas 62% (23/37) reported having an Android phone (9/37, 24%, either provided an invalid response or reported that they did not know what type of phone they owned; the remaining participant had missing data). Overall, 76% (29/38) of participants reported that they used the internet and apps on their phones. Participants were asked how likely they would be to use a smartphone if one was provided for treatment on a scale of 1 (“I would never use it”) to 10 (“I would definitely use it”); the average rating was 8.42 (SD 2.82).

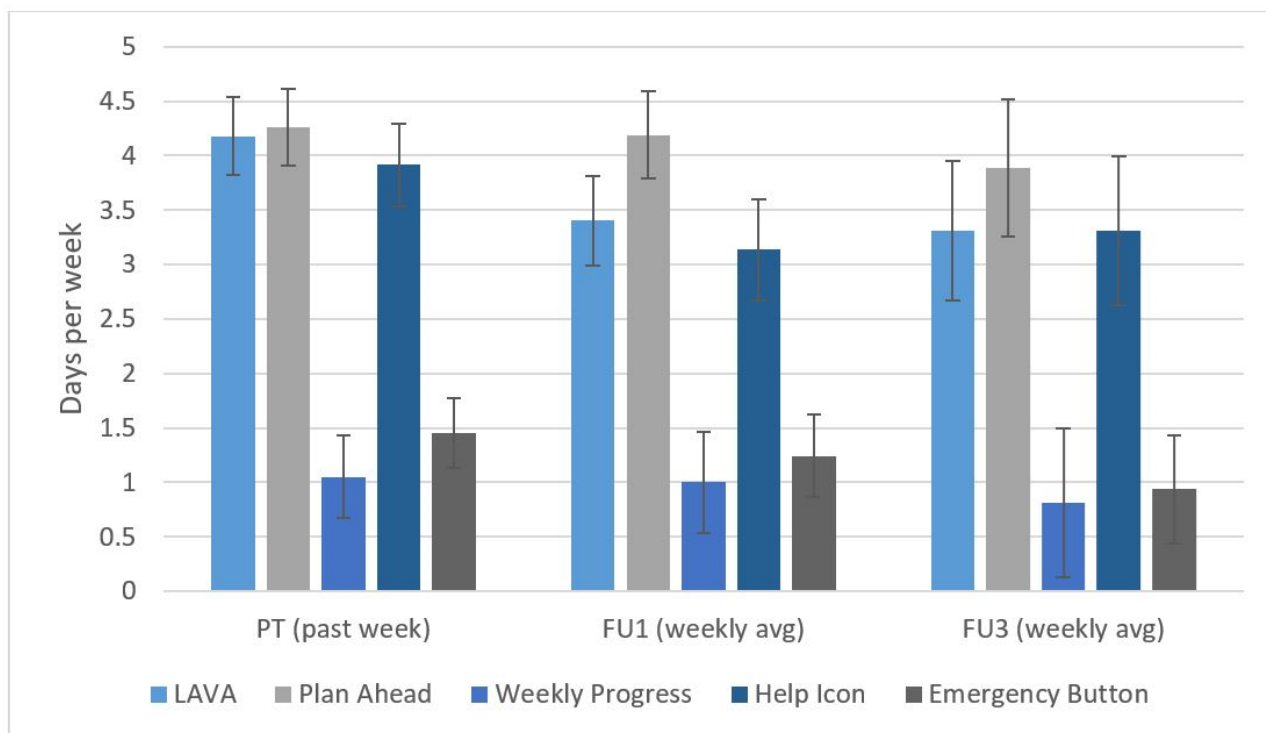
Self-reported App Engagement

Self-reported app use data were obtained at posttreatment and FU assessments for 96% (54/56) of participants. Of these, 72% (39/54) reported any app use at posttreatment, 69% (37/54) reported any app use at FU1, and 37% (20/54) reported any app use at FU3. Chi-square tests indicated that the proportion of participants reporting any app use at posttreatment was significantly greater than the proportion reporting app use at FU3 ($\chi^2_{2} \geq 20.3$; $P < .001$); significant decreases were also identified between FU1 and FU3 ($\chi^2_{2} \geq 11.5$; $P = .001$). Considering only those participants who reported app use at each time point, 54% (21/39) reported using the app at least three times per week at posttreatment. Of those who used the app fewer than 3 times per week, common reasons included forgetting (10/18, 56%) and that it was difficult to use (4/18, 22%). Participants with any app use at posttreatment reported using the LAVA library, Plan Ahead, and Weekly Progress features an average of 4.18 (SD 2.19), 4.26 (SD 2.19), and 3.92

(SD 2.37) days per week, respectively; by the 3-month FU (FU3), participants reported using these features 3.31 (SD 2.57), 3.89 (SD 2.68), and 3.31 (SD 2.73) days per week on average (Figure 2). The Emergency button was used 1.05 (SD 1.88) and

0.81 (SD 1.80) days per week at posttreatment and FU3, respectively, and the Help page was used 1.44 (SD 1.98) and 0.94 (SD 1.98) days per week at posttreatment and FU3, respectively.

Figure 2. Weekly app use by component. FU: follow-up; LAVA: Life Areas, Values, and Activities; PT: posttreatment.



App Component Usefulness

For each app component, participants rated their agreement with the statement that the app component was *a useful part of treatment* on a scale of 1 (strongly disagree) to 5 (strongly agree). The average ratings across all app components indicated that participants generally agreed that each component was useful (LAVA library: mean 4.39, SD 0.72; Plan Ahead: mean 4.33, SD 0.63; Emergency button: mean 3.97, SD 1.00; Weekly Progress: mean 4.22, SD 0.89; Help icon mean: 3.94, SD 0.87). Two-tailed paired *t* tests comparing usefulness ratings between components indicated that ratings for the LAVA library were significantly higher on average than the Emergency button ($t_{40}=2.56$, SE 0.13; $P=.01$) and Help Page ($t_{40}=3.59$, SE 0.87; $P=.001$), and ratings for the Plan Ahead feature were significantly higher than the Help Page ($t_{39}=2.69$, SE 0.76; $P=.01$).

Participants were also asked to provide open-ended feedback about the smartphone-enhanced treatment, including suggestions for improvement. Of the 51 qualitative responses, 73% (37) expressed purely positive feedback. Written comments cited the usefulness of the treatment and highlighted the novelty of the smartphone-enhanced intervention and its utility in facilitating activity planning, helping participants to “be

consistent”, and supporting their recovery. Positive comments included, “It was a good way to get me thinking about how my activities affect my emotions,” “[The app] keeps me on track to be more consistent on a daily basis with responsibilities,” and “It gave me more recovery tools to work with.” Suggestions for improvement included increasing overall ease of use, “more scheduling options” for activities, and providing a smartphone for use after the study FU period.

Reasons for Not Using App Components

Across all components except for the Help page, forgetting to use the app component was by far the most frequently endorsed reason for lack of use (activity scheduling: 19/56, 34%; LAVA: 10/56, 18%; Emergency: 9/56, 16%; and Weekly Progress: 10/56, 18%), whereas not having the smartphone when the participant needed to use the app was generally the second-most endorsed (activity scheduling: 5/56, 9%; LAVA: 6/56, 11%; Emergency: 5/56, 9%; and Weekly Progress: 5/56, 9%). For the Help page, the most frequently endorsed reason was lack of need for the feature (19/56, 34%), whereas forgetting and not having a smartphone were the second-most endorsed reasons (3/56 each, 5%). To provide an example of the distribution of responses for 1 main component, Table 2 shows the reasons endorsed by participants for not scheduling activities.

Table 2. Reasons for not scheduling activities (n=56).

If there were days when you did NOT have an activity scheduled, it was because: (check all that apply)	Frequency, n (%)
This does not apply to me because I scheduled an activity on most days.	18 (32)
I did not remember to use the Daily Plan.	19 (34)
I did not have the smartphone with me when I needed to fill it out.	5 (9)
Filling it out took too much time or effort.	3 (5)
I had technical difficulties with the smartphone.	3 (5)
I did not think it would be helpful to me or my treatment goals.	1 (2)
Filling it out made me uncomfortable.	1 (2)
It was difficult to understand how to use it.	2 (4)
Other	6 (11)
Lost or do not have phone	1 (2)
Incarcerated or hospitalized	2 (4)
Other or undisclosed	3 (5)
Missing or no response	13 (23)

Discussion

Self-reported App Use

With regard to self-reported app use, the majority (37/54, 69%) of participants reported continuing to use their LETS ACT app until 1 month posttreatment, but this proportion decreased significantly (to 20/54, 37%) by 3 months posttreatment. Research suggests that it is typical for mobile app use to quickly drop off. Indeed, many app users use an app only once; on average, user retention is 42% after 1 month and 27% after 3 months [31]. It appears, then, that the LETS ACT app follows a similar rate of decline in use to apps more broadly, although the level of retention was higher than average at both 1 and 3 months posttreatment. This level of attrition and lower continual engagement follows typical patterns of app use in mobile health interventions for SUD [32,33].

Regarding specific app components, participants generally agreed that each app component was a useful part of their treatment. The LAVA library and Plan Ahead feature, which were both essential to the core homework of activity planning, were the most used components and were also rated as more useful compared with features such as the Emergency button and Help page.

Challenges and Recommendations Regarding Integration of Smartphone Technology

In this study, when participants did not use their apps or the individual app components, they generally reported that this was because of either forgetting or not having the smartphone with them. Both reasons for nonuse may reflect the drawbacks of giving participants a study smartphone rather than downloading the app on their own phones. This highlights a critical decision in smartphone-enhanced intervention research, that is, whether to provide smartphones for each participant to ensure consistent phone access or to offer an app that participants can download on their own phones to prioritize utility, ease of use, and generalizability. Research suggests that

low-income people who use drugs have high rates of smartphone ownership, but that they tend to cycle through smartphones and have inconsistent access to wireless data [23], which presents a challenge for assessing the effectiveness of a smartphone-enhanced intervention. This study opted to provide smartphones to participants but found that participants often forgot to use them, which may be in part because they were not using the study phone as their primary phone. Giving participants the option to *either* download the app on their own phone *or* use a study smartphone may ultimately be ideal, although it requires additional resources (ie, developing app versions for both Android and iOS smartphones).

Forgetting to use the app was very common, so future studies examining the integration of smartphone apps into treatment may consider a range of strategies to mitigate this. This could include adding app features that target engagement. For example, reminders in the form of push notifications have been shown to significantly increase app use in smartphone interventions [34,35]. In-app direct messaging may also have the potential to increase engagement, as it provides a means for patients to communicate directly with treatment providers. As an additional strategy, therapists in smartphone-enhanced interventions could use time during treatment sessions to specifically target out-of-session app engagement with the ultimate goal of increasing homework compliance. For example, therapists could help participants make a plan to open their treatment apps at a regular time each day and to link this behavior to other daily activities (eg, “when I get out of bed each morning, I will open my app to see what activities I have planned for the day”).

Participants reporting that they did not have their smartphones with them when needed may reflect challenges inherent to maintaining treatment engagement among low-income people with SUD, many of whom experience significant instability in their daily lives. Some participants were incarcerated during the FU period, whereas others had smartphones that were lost or stolen; still others reported in informal conversations with

research staff that they were nervous about losing their smartphones and chose to store them in a safe place, making them inconvenient to use regularly. Future studies with this population may wish to purchase low-cost smartphones (typically Android devices), which would increase the feasibility of offering replacement devices when needed. Given that this study found that most participants owned Android devices (vs iPhones [Apple Inc]) at pretreatment, they may have the benefit of being familiar to participants in addition to being more affordable.

The results of this study must be interpreted in the context of its limitations. Reliance on self-reported measures of app engagement may be associated with recall bias. In addition, the study sample was recruited from an intensive outpatient treatment center serving a primarily low-income, high school-educated clientele, and participation in the current intervention was offered in addition to standard treatment; the results may not be generalizable to other populations or treatment settings.

Conclusions

There is a clear need for evidence-based SUD treatments that can be delivered at a low cost, and it is essential to find new

and effective ways to engage participants in these treatments. Despite notable growth in the area of app-based psychological interventions, very little research has examined the impact of introducing smartphones into in-person therapy, especially in the area of SUD treatment. This study found evidence that integrating a smartphone app into a BA treatment for SUD is feasible and well accepted and that participants found the app useful for planning value-based activities, which is the core task of BA. However, the study also found that engagement with the app decreased over the FU period and that participants frequently reported forgetting to use the app, highlighting the need for further efforts to sustain out-of-session engagement over time. These findings must be interpreted in light of the specific study methods (eg, the provision of study smartphones to all participants); future research is needed to examine differences in app use when allowing participants to download treatment apps on their own smartphones. Studies are also needed to examine specific treatment contexts and participant characteristics that may be associated with receiving more benefit from smartphone-enhanced interventions.

Conflicts of Interest

None declared.

References

1. 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. 2019. URL: <https://www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUHNationalFindingsReport2018/NSDUHNationalFindingsReport2018.pdf> [accessed 2021-09-09]
2. Olmstead TA, Abraham AJ, Martino S, Roman PM. Counselor training in several evidence-based psychosocial addiction treatments in private US substance abuse treatment centers. *Drug Alcohol Depend* 2012 Jan 1;120(1-3):149-154 [FREE Full text] [doi: [10.1016/j.drugalcdep.2011.07.017](https://doi.org/10.1016/j.drugalcdep.2011.07.017)] [Medline: [21831536](https://pubmed.ncbi.nlm.nih.gov/21831536/)]
3. Wendt DC, Gone JP. Group therapy for substance use disorders: a survey of clinician practices. *J Groups Addict Recover* 2017;12(4):243-259 [FREE Full text] [doi: [10.1080/1556035X.2017.1348280](https://doi.org/10.1080/1556035X.2017.1348280)] [Medline: [30546274](https://pubmed.ncbi.nlm.nih.gov/30546274/)]
4. Roman P, Johnson J. National treatment center study summary report: public treatment centers. Institute for Behavioral Research. 2004. URL: <https://library.net/document/z3orgmez-national-treatment-center-study-summary-report.html> [accessed 2021-09-07]
5. McLellan AT, Lewis DC, O'Brien CP, Kleber HD. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. *JAMA* 2000 Oct 04;284(13):1689-1695. [doi: [10.1001/jama.284.13.1689](https://doi.org/10.1001/jama.284.13.1689)] [Medline: [11015800](https://pubmed.ncbi.nlm.nih.gov/11015800/)]
6. White WL, Boyle M, Loveland D. Alcoholism/addiction as a chronic disease: from rhetoric to clinical reality. *Alcohol Treat Q* 2002 Jul;20(3-4):107-130. [doi: [10.1300/j020v20n03_06](https://doi.org/10.1300/j020v20n03_06)]
7. McKay JR. Continuing care research: what we have learned and where we are going. *J Subst Abuse Treat* 2009 Mar;36(2):131-145 [FREE Full text] [doi: [10.1016/j.jsat.2008.10.004](https://doi.org/10.1016/j.jsat.2008.10.004)] [Medline: [19161894](https://pubmed.ncbi.nlm.nih.gov/19161894/)]
8. Hubbard RL, Craddock SG, Anderson J. Overview of 5-year followup outcomes in the drug abuse treatment outcome studies (DATOS). *J Subst Abuse Treat* 2003 Oct;25(3):125-134. [doi: [10.1016/s0740-5472\(03\)00130-2](https://doi.org/10.1016/s0740-5472(03)00130-2)] [Medline: [14670518](https://pubmed.ncbi.nlm.nih.gov/14670518/)]
9. Daughters SB, Magidson JF, Lejuez C, Chen Y. LETS ACT: a behavioral activation treatment for substance use and depression. *Adv Dual Diagn* 2016 Jul 2;9(2/3):74-84. [doi: [10.1108/add-02-2016-0006](https://doi.org/10.1108/add-02-2016-0006)]
10. Murphy JG, Correia CJ, Barnett NP. Behavioral economic approaches to reduce college student drinking. *Addict Behav* 2007 Nov;32(11):2573-2585. [doi: [10.1016/j.addbeh.2007.05.015](https://doi.org/10.1016/j.addbeh.2007.05.015)] [Medline: [17600631](https://pubmed.ncbi.nlm.nih.gov/17600631/)]
11. Ekers D, Godfrey C, Gilbody S, Parrott S, Richards DA, Hammond D, et al. Cost utility of behavioural activation delivered by the non-specialist. *Br J Psychiatry* 2011 Dec;199(6):510-511. [doi: [10.1192/bjp.bp.110.090266](https://doi.org/10.1192/bjp.bp.110.090266)] [Medline: [21947655](https://pubmed.ncbi.nlm.nih.gov/21947655/)]

12. Daughters SB, Braun AR, Sargeant MN, Reynolds EK, Hopko DR, Blanco C, et al. Effectiveness of a brief behavioral treatment for inner-city illicit drug users with elevated depressive symptoms: the life enhancement treatment for substance use (LETS Act!). *J Clin Psychiatry* 2008 Jan;69(1):122-129. [doi: [10.4088/jcp.v69n0116](https://doi.org/10.4088/jcp.v69n0116)] [Medline: [18312046](https://pubmed.ncbi.nlm.nih.gov/18312046/)]
13. Magidson JF, Gorka SM, MacPherson L, Hopko DR, Blanco C, Lejuez CW, et al. Examining the effect of the Life Enhancement Treatment for Substance Use (LETS ACT) on residential substance abuse treatment retention. *Addict Behav* 2011 Jun;36(6):615-623 [FREE Full text] [doi: [10.1016/j.addbeh.2011.01.016](https://doi.org/10.1016/j.addbeh.2011.01.016)] [Medline: [21310539](https://pubmed.ncbi.nlm.nih.gov/21310539/)]
14. Daughters SB, Magidson JF, Anand D, Seitz-Brown CJ, Chen Y, Baker S. The effect of a behavioral activation treatment for substance use on post-treatment abstinence: a randomized controlled trial. *Addiction* 2018 Mar;113(3):535-544 [FREE Full text] [doi: [10.1111/add.14049](https://doi.org/10.1111/add.14049)] [Medline: [28963853](https://pubmed.ncbi.nlm.nih.gov/28963853/)]
15. Mausbach BT, Moore R, Roesch S, Cardenas V, Patterson TL. The relationship between homework compliance and therapy outcomes: an updated meta-analysis. *Cognit Ther Res* 2010 Oct;34(5):429-438 [FREE Full text] [doi: [10.1007/s10608-010-9297-z](https://doi.org/10.1007/s10608-010-9297-z)] [Medline: [20930925](https://pubmed.ncbi.nlm.nih.gov/20930925/)]
16. Lebeau RT, Davies CD, Culver NC, Craske MG. Homework compliance counts in cognitive-behavioral therapy. *Cogn Behav Ther* 2013;42(3):171-179. [doi: [10.1080/16506073.2013.763286](https://doi.org/10.1080/16506073.2013.763286)] [Medline: [23419077](https://pubmed.ncbi.nlm.nih.gov/23419077/)]
17. Rees CS, McEvoy P, Nathan PR. Relationship between homework completion and outcome in cognitive behaviour therapy. *Cogn Behav Ther* 2005;34(4):242-247. [doi: [10.1080/16506070510011548](https://doi.org/10.1080/16506070510011548)] [Medline: [16319035](https://pubmed.ncbi.nlm.nih.gov/16319035/)]
18. Savic M, Best D, Rodda S, Lubman DI. Exploring the focus and experiences of smartphone applications for addiction recovery. *J Addict Dis* 2013;32(3):310-319. [doi: [10.1080/10550887.2013.824331](https://doi.org/10.1080/10550887.2013.824331)] [Medline: [24074196](https://pubmed.ncbi.nlm.nih.gov/24074196/)]
19. Ly KH, Janni E, Wrede R, Sedem M, Donker T, Carlbring P, et al. Experiences of a guided smartphone-based behavioral activation therapy for depression: a qualitative study. *Internet Interv* 2015 Mar;2(1):60-68. [doi: [10.1016/j.invent.2014.12.002](https://doi.org/10.1016/j.invent.2014.12.002)]
20. Luxton DD, Hansen RN, Stanfill K. Mobile app self-care versus in-office care for stress reduction: a cost minimization analysis. *J Telemed Telecare* 2014 Dec;20(8):431-435. [doi: [10.1177/1357633X14555616](https://doi.org/10.1177/1357633X14555616)] [Medline: [25316037](https://pubmed.ncbi.nlm.nih.gov/25316037/)]
21. Mobile fact sheet. Pew Research Center. 2019. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2021-09-07]
22. Dahne J, Lejuez CW. Smartphone and mobile application utilization prior to and following treatment among individuals enrolled in residential substance use treatment. *J Subst Abuse Treat* 2015 Nov;58:95-99 [FREE Full text] [doi: [10.1016/j.jsat.2015.06.017](https://doi.org/10.1016/j.jsat.2015.06.017)] [Medline: [26231698](https://pubmed.ncbi.nlm.nih.gov/26231698/)]
23. Tofighi B, Leonard N, Greco P, Hadavand A, Acosta MC, Lee JD. Technology use patterns among patients enrolled in inpatient detoxification treatment. *J Addict Med* 2019;13(4):279-286 [FREE Full text] [doi: [10.1097/ADM.0000000000000494](https://doi.org/10.1097/ADM.0000000000000494)] [Medline: [30589653](https://pubmed.ncbi.nlm.nih.gov/30589653/)]
24. Carreiro S, Newcomb M, Leach R, Ostrowski S, Boudreaux ED, Amante D. Current reporting of usability and impact of mHealth interventions for substance use disorder: a systematic review. *Drug Alcohol Depend* 2020 Oct 1;215:108201. [doi: [10.1016/j.drugalcdep.2020.108201](https://doi.org/10.1016/j.drugalcdep.2020.108201)] [Medline: [32777691](https://pubmed.ncbi.nlm.nih.gov/32777691/)]
25. Fowler LA, Holt SL, Joshi D. Mobile technology-based interventions for adult users of alcohol: a systematic review of the literature. *Addict Behav* 2016 Nov;62:25-34. [doi: [10.1016/j.addbeh.2016.06.008](https://doi.org/10.1016/j.addbeh.2016.06.008)] [Medline: [27310031](https://pubmed.ncbi.nlm.nih.gov/27310031/)]
26. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone apps targeting alcohol and illicit substance use: systematic search in commercial app stores and critical content analysis. *JMIR Mhealth Uhealth* 2019 Apr 22;7(4):e11831 [FREE Full text] [doi: [10.2196/11831](https://doi.org/10.2196/11831)] [Medline: [31008713](https://pubmed.ncbi.nlm.nih.gov/31008713/)]
27. Tønning ML, Kessing LV, Bardram JE, Faurholt-Jepsen M. Methodological challenges in randomized controlled trials on smartphone-based treatment in psychiatry: systematic review. *J Med Internet Res* 2019 Oct 27;21(10):e15362 [FREE Full text] [doi: [10.2196/15362](https://doi.org/10.2196/15362)] [Medline: [31663859](https://pubmed.ncbi.nlm.nih.gov/31663859/)]
28. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. *JAMA Psychiatry* 2014 May;71(5):566-572 [FREE Full text] [doi: [10.1001/jamapsychiatry.2013.4642](https://doi.org/10.1001/jamapsychiatry.2013.4642)] [Medline: [24671165](https://pubmed.ncbi.nlm.nih.gov/24671165/)]
29. Rawson RA, Shoptaw SJ, Obert JL, McCann MJ, Hasson AL, Marinelli-Casey PJ, et al. An intensive outpatient approach for cocaine abuse treatment. The Matrix model. *J Subst Abuse Treat* 1995;12(2):117-127. [doi: [10.1016/0740-5472\(94\)00080-b](https://doi.org/10.1016/0740-5472(94)00080-b)] [Medline: [7623389](https://pubmed.ncbi.nlm.nih.gov/7623389/)]
30. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22-33;quiz 34. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
31. 2019 mobile app benchmark report to inform your 2020 strategy. Localytics. URL: <https://uplandsoftware.com/localytics/resources/cheat-sheet/2019-mobile-app-benchmark-report-to-inform-your-2020-strategy/> [accessed 2021-03-31]
32. Kazemi DM, Borsari B, Levine MJ, Li S, Lamberson KA, Matta LA. A systematic review of the mHealth interventions to prevent alcohol and substance abuse. *J Health Commun* 2017 May;22(5):413-432 [FREE Full text] [doi: [10.1080/10810730.2017.1303556](https://doi.org/10.1080/10810730.2017.1303556)] [Medline: [28394729](https://pubmed.ncbi.nlm.nih.gov/28394729/)]
33. Attwood S, Parke H, Larsen J, Morton KL. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application. *BMC Public Health* 2017 May 17;17(1):394 [FREE Full text] [doi: [10.1186/s12889-017-4358-9](https://doi.org/10.1186/s12889-017-4358-9)] [Medline: [28511698](https://pubmed.ncbi.nlm.nih.gov/28511698/)]

34. Hernández-Reyes A, Cámara-Martos F, Molina Recio G, Molina-Luque R, Romero-Saldaña M, Moreno Rojas R. Push notifications from a mobile app to improve the body composition of overweight or obese women: randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Feb 12;8(2):e13747 [FREE Full text] [doi: [10.2196/13747](https://doi.org/10.2196/13747)] [Medline: [32049065](https://pubmed.ncbi.nlm.nih.gov/32049065/)]
35. Brindal E, Hendrie GA, Freyne J, Noakes M. Incorporating a static versus supportive mobile phone app into a partial meal replacement program with face-to-face support: randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Apr 18;6(4):e41 [FREE Full text] [doi: [10.2196/mhealth.7796](https://doi.org/10.2196/mhealth.7796)] [Medline: [29669704](https://pubmed.ncbi.nlm.nih.gov/29669704/)]

Abbreviations

BA: behavioral activation

DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*

FU: follow-up

LAVA: Life Areas, Values, and Activities

LETS ACT: Life Enhancement Treatment for Substance Use

SUD: substance use disorder

TAU: treatment as usual

Edited by G Eysenbach; submitted 23.11.20; peer-reviewed by Z Huang, J Hidalgo, S Stull; comments to author 05.02.21; revised version received 14.05.21; accepted 01.08.21; published 03.11.21.

Please cite as:

Paquette CE, Rubalcava DT, Chen Y, Anand D, Daughters SB

A Mobile App to Enhance Behavioral Activation Treatment for Substance Use Disorder: App Design, Use, and Integration Into Treatment in the Context of a Randomized Controlled Trial

JMIR Form Res 2021;5(11):e25749

URL: <https://formative.jmir.org/2021/11/e25749>

doi: [10.2196/25749](https://doi.org/10.2196/25749)

PMID: [34730535](https://pubmed.ncbi.nlm.nih.gov/34730535/)

©Catherine E Paquette, Dillon T Rubalcava, Yun Chen, Deepika Anand, Stacey B Daughters. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 03.11.2021. 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 Video Self-Modeling Intervention Using Virtual Reality Plus Physical Practice for Freezing of Gait in Parkinson Disease: Feasibility and Acceptability Study

Lina Goh¹, BAppSc; Natalie E Allen¹, PhD; Naseem Ahmadpour², PhD; Kaylena A Ehgoetz Martens³, PhD; Jooeun Song¹, PhD; Lindy Clemson¹, PhD; Simon J G Lewis⁴, MD; Hamish G MacDougall⁵, PhD; Colleen G Canning¹, PhD

¹Sydney School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Camperdown, Australia

²Sydney School of Architecture, Design and Planning, The University of Sydney, Camperdown, Australia

³Department of Kinesiology, Faculty of Health, The University of Waterloo, Waterloo, ON, Canada

⁴Brain and Mind Centre, The University of Sydney, Camperdown, Australia

⁵School of Psychology, Faculty of Science, The University of Sydney, Camperdown, Australia

Corresponding Author:

Natalie E Allen, PhD

Sydney School of Health Sciences

Faculty of Medicine and Health

The University of Sydney

Level 7 Susan Wakil Health Building D18

Camperdown, 2006

Australia

Phone: 61 9351 9016

Email: natalie.allen@sydney.edu.au

Abstract

Background: Despite optimal medical and surgical intervention, freezing of gait commonly occurs in people with Parkinson disease. Action observation via video self-modeling, combined with physical practice, has potential as a noninvasive intervention to reduce freezing of gait.

Objective: The aim of this study is to determine the feasibility and acceptability of a home-based, personalized video self-modeling intervention delivered via a virtual reality head-mounted display (HMD) to reduce freezing of gait in people with Parkinson disease. The secondary aim is to investigate the potential effect of this intervention on freezing of gait, mobility, and anxiety.

Methods: The study was a single-group pre-post mixed methods pilot trial for which 10 participants with Parkinson disease and freezing of gait were recruited. A physiotherapist assessed the participants in their homes to identify person-specific triggers of freezing and developed individualized movement strategies to overcome freezing of gait. 180° videos of the participants successfully performing their movement strategies were created. Participants watched their videos using a virtual reality HMD, followed by physical practice of their strategies in their own homes over a 6-week intervention period. The primary outcome measures included the feasibility and acceptability of the intervention. Secondary outcome measures included freezing of gait physical tests and questionnaires, including the Timed Up and Go Test, 10-meter walk test, Goal Attainment Scale, and Parkinson Anxiety Scale.

Results: The recruitment rate was 24% (10/42), and the retention rate was 90% (9/10). Adherence to the intervention was high, with participants completing a mean of 84% (SD 49%) for the prescribed video viewing and a mean of 100% (SD 56%) for the prescribed physical practice. One participant used the virtual reality HMD for 1 week and completed the rest of the intervention using a flat-screen device because of a gradual worsening of his motion sickness. No other adverse events occurred during the intervention or assessment. Most of the participants found using the HMD to view their videos interesting and enjoyable and would choose to use this intervention to manage their freezing of gait in the future. Five themes were constructed from the interview data: reflections when seeing myself, my experience of using the virtual reality system, the role of the virtual reality system in supporting my learning, developing a deeper understanding of how to manage my freezing of gait, and the impact of the intervention on my daily activities. Overall, there were minimal changes to the freezing of gait, mobility, or anxiety measures from baseline to postintervention, although there was substantial variability between participants. The intervention showed potential in reducing anxiety in participants with high levels of anxiety.

Conclusions: Video self-modeling using an immersive virtual reality HMD plus physical practice of personalized movement strategies is a feasible and acceptable method of addressing freezing of gait in people with Parkinson disease.

(*JMIR Form Res* 2021;5(11):e28315) doi:[10.2196/28315](https://doi.org/10.2196/28315)

KEYWORDS

Parkinson disease; freezing of gait; action observation; video self-modelling; virtual reality

Introduction

Background

Parkinson disease is a progressive neurological condition that affects approximately 6 million people worldwide [1]. People with Parkinson disease can present with a variety of motor impairments such as tremor, slow movements, gait, and balance disorders, as well as nonmotor impairments such as reduced cognition, depression, anxiety, and sleep disorders [2]. Freezing of gait, defined as a brief, episodic, absence or marked reduction of forward progression of the feet despite the intention to walk, is a complex phenomenon that may also be present in people with Parkinson disease [3,4]. People with freezing of gait often describe their feet as being *glued to the ground*, which can lead to reduced mobility, falls, poor quality of life, and increased health care costs [4-8].

The pathophysiology underlying freezing of gait remains poorly understood, although it is suggested to result from dysfunction in neural networks across motor, affective, and cognitive domains [9-12]. People with freezing of gait are more likely to exhibit decreased gait automaticity and increased gait variability [13], as well as motor fluctuations and dyskinesia [14]. Furthermore, cognitive deficits and anxiety also appear to be more pronounced in people with freezing of gait than without [14-17]. Freezing of gait is also more frequent and severe in conditions of high levels of anxiety compared with low levels [18].

First-line treatments for freezing of gait consist predominantly of pharmacological interventions to maintain a good *on* state [4]. However, freezing of gait can persist despite these regimes. Nonpharmacological interventions are often used in conjunction with pharmacological interventions, with the most common being physiotherapy. Although several reviews have shown that these nonpharmacological interventions are effective, their results have been modest [19-23]. This might be because of the heterogeneity of the interventions, small study sample sizes, and limitations of relying on self-report to assess freezing of gait [21,24].

A recent systematic review performed subgroup analyses to determine which types of physiotherapy interventions may be the most useful in managing freezing of gait. The results showed that action observation training had a statistically significant effect on reducing freezing of gait [21]. In the 4 action observation studies, participants with freezing of gait watched videos of actors perform movement strategies designed to overcome freezing, followed by physical practice of these strategies under the supervision of a physiotherapist [25-28]. Observation of actions performed by others is understood to activate neural structures in the brain that execute the same

actions, thus facilitating motor learning and performance [29]. This is then further reinforced by physical practice.

Although the results from these action observation studies were positive, the strength of this evidence was weak, and its clinical significance was unclear. This was because of the small number of moderate-quality studies included in the meta-analysis and the use of the New Freezing of Gait Questionnaire (NFOG-Q) [30] as the outcome, as this questionnaire was previously shown to be insufficiently responsive to detect small changes in freezing severity [24] and may not be a good indicator of the real symptom burden [31]. Furthermore, the implementation of action observation as an intervention for freezing of gait was limited in the following ways: (1) participants watched actors without Parkinson disease perform movement strategies, which might not be an accurate reflection of motor performance by people with Parkinson disease; (2) videos demonstrated generalized strategies that might not be relevant or appropriate for the individual; (3) videos showed the use of movement strategies to overcome freezing of gait in clinical settings, which might reduce ecological validity as patients typically present with freezing of gait at home [32]; and (4) participants watching videos on flat-screen devices could be vulnerable to distractions in their immediate environment, which might impact motor learning.

Video self-modeling, a form of observational learning that requires the observer to watch and learn from one's own positive behavior, may be useful for people with Parkinson disease [33]. The activation of neural structures previously described is maximized when the observed actions are familiar to the observer and comprise movement that the observer is able to perform [29]. Therefore, we hypothesized that the use of video self-modeling would provide salient cues to improve motor learning and performance, as well as promote self-efficacy by strengthening beliefs in one's ability to overcome freezing of gait [34]. People with freezing of gait may benefit further if they observe themselves using personalized strategies that address their specific motor, affective, and cognitive triggers of freezing. In addition, videos of situations at home that provoke freezing of gait (and strategies to successfully overcome freezing) may be of greater relevance. Furthermore, the use of a virtual reality head-mounted display (HMD) removes distractors in the environment, directing full attention to viewing videos.

Objective

Therefore, the aim of this study is to investigate the feasibility and acceptability of a home-based, personalized video self-modeling intervention delivered via a virtual reality HMD to reduce freezing of gait in people with Parkinson disease. The

secondary aim is to investigate the potential effect of this intervention on freezing of gait, mobility, and anxiety.

Methods

Design

A single-group pre-post mixed methods pilot trial was conducted from April 2019 to April 2020. Ethics approval was obtained from the University of Sydney Human Research Ethics Committee (project number 2018/893), and written informed consent was obtained from all participants. The trial was registered with the Australian New Zealand Clinical Trials Registry (ANZCTR12619000139178).

Participants

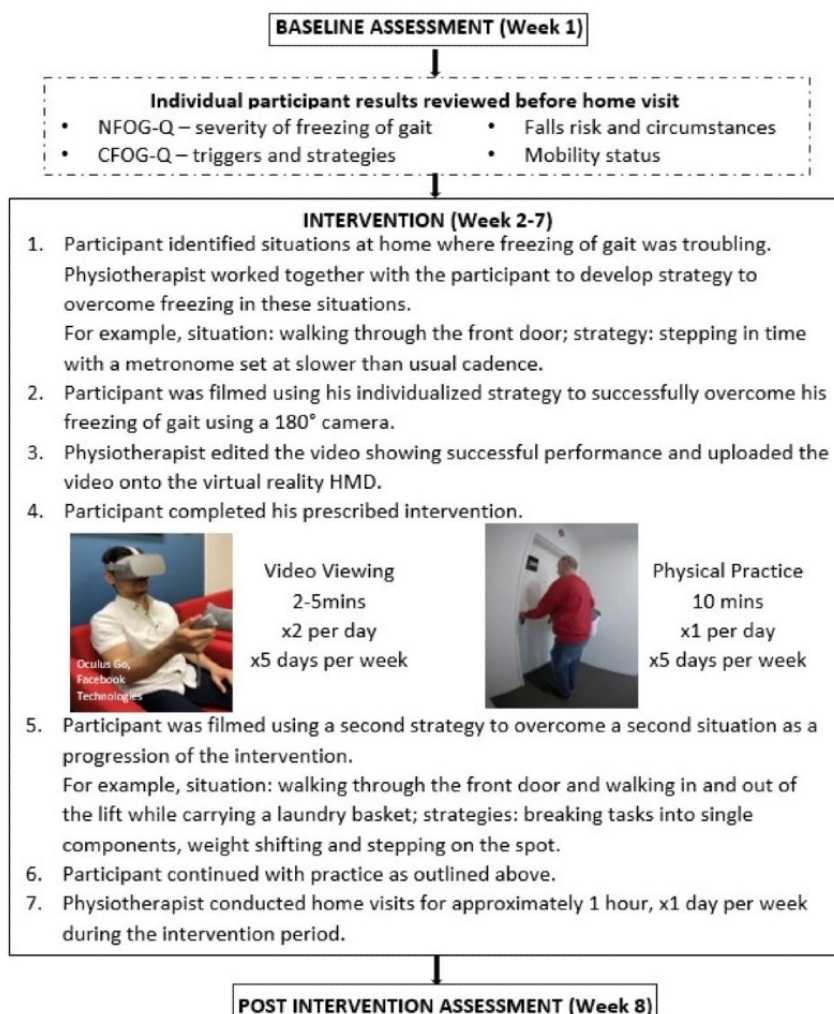
A total of 10 participants were recruited from existing databases of people with Parkinson disease at the University of Sydney and from Parkinson’s New South Wales support groups. The following inclusion criteria were used: (1) diagnosis of idiopathic Parkinson disease, (2) presence of freezing of gait (defined as having a score of ≥ 1 on question 2 and score of ≥ 2 on question 4 of the NFOG-Q), (3) stable dopaminergic

medication regime for at least 4 weeks before commencing the study, (4) ability to walk independently with or without a walking aid, and (5) living in the greater Sydney metropolitan area. Participants were excluded if they had (1) any medical conditions that would interfere with the study safety and conduct, such as unstable cardiovascular disease and neurological conditions other than Parkinson disease; (2) cognitive impairment defined as having a score of < 24 on the Mini Mental State Examination [35]; (3) newly commenced deep brain stimulation or changes in stimulation parameters within 6 months before participating in the study; and (4) significant head tremor or motion sickness limiting the ability to use a virtual reality HMD.

Intervention

The intervention protocol (including printed information and instructions to the participants) is described in detail in [Multimedia Appendix 1](#) [23,25,36-38] and illustrated in [Figure 1](#). In brief, a physiotherapist (LG) delivered 6-8 home visits over 6 weeks, with each visit lasting approximately 60 minutes. The interventions were delivered, and practice was completed when participants were in their *on* phase, that is, when their medications were working optimally.

Figure 1. Example of the intervention. CFOG-Q: Characterizing Freezing of Gait Questionnaire; HMD: head-mounted display; NFOG-Q: New Freezing of Gait Questionnaire.



Before the first home visit, the physiotherapist reviewed the participants' medical history, including disease and freezing of gait severity (NFOG-Q), mobility status, and the circumstances of any falls. The physiotherapist also evaluated the baseline results of participants' Characterizing Freezing of Gait Questionnaire (CFOG-Q) to determine their personal motor, cognitive, emotional, and environmental triggers of freezing of gait, as well as any strategies previously used to overcome freezing [36].

During the first home visit, participants identified a situation at home, where freezing of gait was troubling. Examples of situations included turning around in tight spaces and walking through doorways while performing an additional motor task. The physiotherapist assessed the participants' performance in this situation and worked with participants to develop a suitable personalized strategy to help them overcome their freezing. Examples of strategies included stepping in time to the beats of an external rhythmic auditory cueing device, such as a metronome typically set at lower than usual cadence; self-initiated movement strategies, such as counting, shifting weight from side to side, and/or simplifying complex tasks; and progressing by increasing the complexity of tasks where appropriate.

Once a situation was identified and a strategy developed and practiced by the participants, a 3D camera (Mirage, Lenovo) was used to produce 180° videos showing the participants using their personalized strategy to successfully overcome freezing of gait in the situation identified. The videos were edited to show the participants' successful performance 3 times and lasted 2-5 minutes. The videos were created to be compatible with a virtual reality HMD (Oculus Go, Facebook Technologies). Participants were instructed on the use of the HMD and asked to watch their personal video using the HMD while seated twice a day, 5 days per week for 6 weeks. After one of the 2 daily video viewings, participants performed the physical practice of their strategy, as shown in their video for 10 minutes. Participants were also provided with a printed guide on how to use and navigate the virtual reality system.

The physiotherapist monitored participants' progress during home visits. Once participants were able to consistently use their strategy to overcome freezing of gait in the first situation, a second situation was identified, and a second strategy was developed. A second video was created for each participant. This second video involved the participant performing the same task in a different environment, the same task with increased complexity, or a different task altogether. Strategies used to overcome freezing of gait were adjusted as appropriate depending on the situation, the person-specific freezing of gait triggers, and preferences. Participants were then asked to watch their second video and perform its associated physical practice 4 times a week, and the first video and its associated physical practice once a week. If participants continued to make further progress, a third situation was introduced, and this process continued for the duration of the intervention period. The introduction of additional videos was determined by the physiotherapist and was based on clinical judgment and participants' self-assessment of their progress.

Outcome Measures

The primary outcome measures assessed the feasibility and acceptability of the intervention. Measures of feasibility included recruitment rate, retention rate, and adherence to the intervention (by recording the number of daily video viewings and physical practice) using self-report logbooks and adverse events associated with the intervention. Measures of acceptability included a modified Players Experience of Need Satisfaction (PENS) Questionnaire [39] and a semistructured interview at postintervention. In the modified PENS Questionnaire (Multimedia Appendix 2 [39]), participants were asked to reflect on their experience of using the virtual reality system, which comprised the HMD, the handheld control, and the experience of navigating and viewing the videos. Participants rated their interest or enjoyment, sense of presence or immersion, competence, and intuitiveness of the controls and rated 2 additional items that were added to the PENS. These were the presence of motion sickness and whether they would use the virtual reality system for the management of their freezing of gait in the future if it was available. The questionnaire included 18 questions across the categories listed above, with each question scored on a 7-point Likert scale, where 1=strongly disagree and 7=strongly agree (higher score is better). In the interviews, participants were asked to describe their experiences of the intervention, including both video viewing and physical practice (Multimedia Appendix 3).

Secondary outcome measures were collected to assess any potential effects of the intervention on freezing of gait, mobility, and anxiety. These measures were collected at baseline and postintervention. Participants also completed the Goal Attainment Scale, with goals in relation to managing freezing of gait at home set at baseline and goal attainment evaluated at postintervention [40].

To obtain freezing of gait measures, participants were videotaped by performing 2 freezing of gait provoking tests: the Ziegler test [41] and the turn-in-place test [42]. In the Ziegler test, participants began in a seated position 3.4 m from a closed door. Participants were asked to stand up, walk forward 1 m to a square outlined with tape on the ground (40×40 cm), perform two 360° turns (clockwise and counter-clockwise) within the square, and walk forward a further 2 m to open the door and walk through the doorway, before returning to sit in the chair. Participants were asked to perform the test as fluently as possible under 3 conditions in the following order: (1) no additional task, (2) with an additional motor task (ie, carrying a tray with a cup of water), and (3) with additional motor and cognitive tasks (ie, carrying a tray with a cup of water and counting backwards by 7 from 100). In the turn-in-place test, participants were asked to turn 360° on the spot, alternating right and left at a self-selected pace for 1 minute.

Two assessors (KAEM and JS) determined the percentage of time frozen for each freezing of gait provoking test, plus the time taken to complete the Ziegler test, via offline video analyses [43,44]. They were blinded to the baseline and postintervention testing conditions. A detailed protocol is described in Multimedia Appendix 4 [4,43]. Interrater reliability between the assessors was excellent. The intraclass correlation

coefficients (two-way mixed effects, absolute agreement) were as follows: Ziegler test percent time frozen=0.962 (from analyses of 29 videos), Ziegler test duration=1.000 (from analyses of 29 videos), and turn-in-place test percent time frozen=0.984 (from analyses of 20 videos). The videos used to determine intraclass correlation coefficients were from baseline *on* and *off* performances of 5 participants where postintervention measures were not available.

Participants also completed the NFOG-Q, which reports the severity and impact of freezing of gait (range 0-28), and the CFOG-Q, where section 2 reports the frequency of freezing of gait triggers (range 0-48). Lower scores indicate less severe freezing of gait in both measures.

The mobility measures included comfortable walking speed (measured over 10 m) [45] and the Timed Up and Go Test [46] in single- and dual-task conditions (ie, counting backward from 100 by 3), with lower scores indicating better mobility. Anxiety was measured using the Parkinson Anxiety Scale (PAS; range 0-48), with lower scores indicating lower levels of anxiety [47].

Participants were assessed at baseline within a week before the start of the intervention and postintervention within a week of completing the intervention. The following demographic information was also collected at baseline: age, gender, severity of Parkinson disease using the Movement Disorder Society—Unified Parkinson's Disease Rating Scale Section III [48] and the Hoehn and Yahr stage [49], cognitive function using the Trail Making Tests A and B [50], and current medication regimen. Assessments were conducted at a university laboratory when participants were in their *on* phase. The freezing of gait provoking tests (both Ziegler and turn-in-place tests) were also repeated in participants' homes when they were in their *off* phase after 12 hours withdrawal of their levodopa medication overnight.

Adverse Events

The presence of motion sickness experienced while using the virtual reality HMD and any falls, injuries, and fatigue during

the intervention and assessments were monitored. Participants were asked to record any adverse events in a logbook and report the events to the researchers. At each home visit, the researcher questioned the participants about the occurrence of any adverse events.

Statistical Analyses

Descriptive statistical analyses of feasibility and secondary outcome measures were conducted using SPSS Statistics for Windows, version 26.0 (IBM Corporation). Interview data were audio-recorded and transcribed verbatim by independent transcribers who were external to the study. NVivo 12 software (version 2, QSR International) was used to code the interview data, and inductive thematic analysis was used to interpret the results [51]. One researcher coded all of the interviews (LG) and 2 other researchers coded parts of the interviews (CGC and NA) such that all data were coded independently by at least 2 researchers. The codes were then compared and grouped to form the main themes through an iterative process. Any differences were discussed in depth until a consensus was reached among the 3 researchers.

Results

Overview

Individual and aggregate participant background information is presented in Table 1. Overall, participants (9 males and 1 female) had a mean age of 70.6 years (SD 7.7 years), were diagnosed with Parkinson disease for an average of 13.3 years (SD 5.2 years), and had moderate to severe disease severity with a mean Movement Disorder Society—Unified Parkinson's Disease Rating Scale Section III score of 37.3 (SD 13.3) [52]. All participants had moderate to severe freezing of gait (NFOG-Q range 10-24/28), and 5 participants had significant anxiety (PAS>14/48). A total of 6 participants were considered recurrent fallers (defined as having more than 2 falls in the past 12 months), with 2 participants experiencing particularly high rates of falls. A total of 2 participants had not fallen in the past 12 months (Table 1).

Table 1. Baseline characteristics of participants.

	Age (years)	Parkinson disease duration (years)	Movement Disorder Society—Unified Parkinson's Disease Rating Scale Section III (0-132) ^a	Hoehn and Yahr scale (1-5) ^a	Mini Mental State Examination (0-30) ^b	Trail Making Test part A (seconds) ^a	Trail Making Test part B (seconds) ^a	New Freezing of Gait Questionnaire (0-28) ^a	Parkinson Anxiety Scale (0-48) ^a	Timed Up and Go Test (seconds) ^a	Number of falls in the past year	Levodopa equivalent daily dose (mg)
P1	81	12	53	3	26	79.0	351.0	22	14	27.0	4	1120
P2	62	14	18	2	30	14.9	26.9	24	9	10.4	1	2083 ^c
P3 ^d	61	18	40	3	30	18.9	49.9	22	10	12.9	6	724
P4	78	11	29	2	30	49.0	169.0	15	10	12.3	6	1300
P5	76	5	45	2	29	13.0	130.2	11	11	12.9	1	600
P6	75	8	27	2	28	59.0	106.0	17	18	13.9	0	900
P7	66	16	35	3	25	41.4	284.5	23	29	15.3	Approximately 550 ^e	600 ^f
P8	60	10	47	2	29	62.0	87.0	24	40	14.5	6	1048
P9 ^g	72	23	22	3	28	28.6	95.0	21	3	9.4	Approximately 230 ^h	1694
P10 ^g	75	16	57	2	28	49.4	149.9	21	20	14.0	0	400
Value, mean (SD)	70.6 (7.7)	13.3 (5.2)	37.3 (13.3)	2.4 (0.5)	28.3 (1.7)	41.5 (22.2)	144.9 (101.7)	20.0 (4.3)	16.4 (11.0)	14.3 (4.8)	N/A ⁱ	1046.9 (527.6)

^aLower scores are better.

^bHigher scores are better.

^cThis participant is on Duodopa therapy.

^dThis participant withdrew because of medical reasons unrelated to the trial.

^eThis participant reported 1 to 2 falls per day.

^fThis participant received deep brain stimulation.

^gThis participant did not receive the complete intervention because of COVID-19 lockdown.

^hThis participant reported 4 to 5 falls per week.

ⁱN/A: not applicable.

Primary Outcome Measures

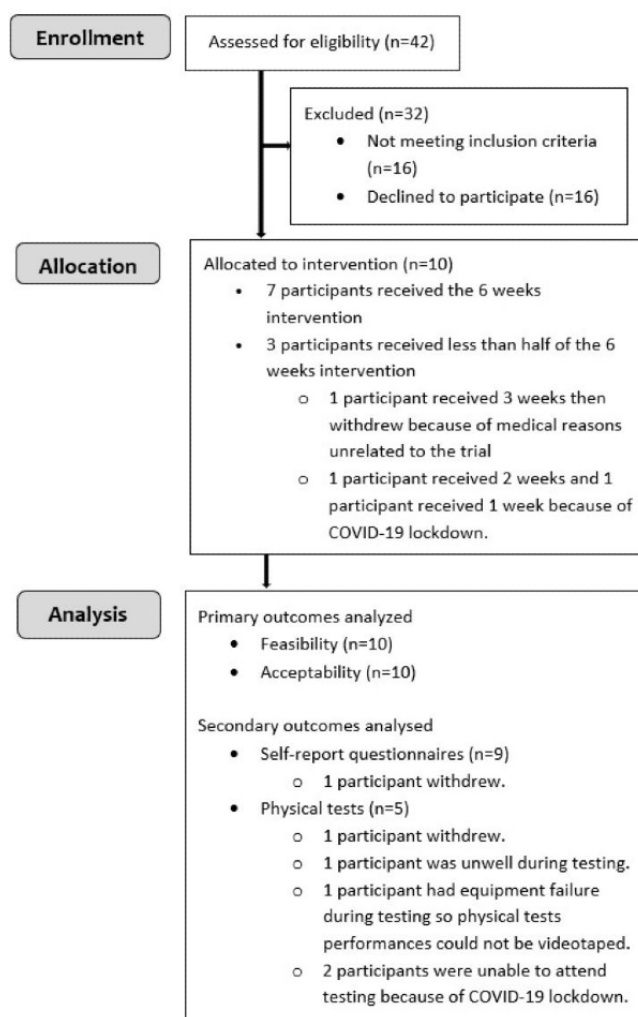
Feasibility

The flow of the participants in this study is shown in [Figure 2](#). A total of 42 potential participants were identified and assessed for eligibility. Of these 42 participants, 10 participants consented to participate, resulting in a 24% (10/42) recruitment rate. A total of 1 participant withdrew after receiving 3 weeks of intervention because of medical reasons unrelated to the study, resulting in a 90% (9/10) retention rate.

All primary and secondary outcome measures were obtained from 10 participants at baseline. At postintervention, all primary outcome measures were obtained, but only some secondary outcome measures were available. A total of 9 participants completed the subjective questionnaires, and 5 completed all the in-person physical tests. In addition to the participant who

withdrew from the study, 1 participant became unwell during testing; 1 participant had equipment failure during testing, so physical test performances could not be videotaped; and 2 participants were unable to attend in-person testing because of a city-wide lockdown from the COVID-19 pandemic.

Adherence to the intervention was good, with participants completing a mean of 84% (SD 49%; range 8%-153%) of the prescribed video viewing and a mean of 100% (SD 56%; range 17%-187%) of the prescribed physical practice ([Table 2](#)). A total of 2 participants received less than 2 weeks of supervised home visits as a result of the COVID-19 lockdown. A total of 1 participant was unable to tolerate the use of the virtual reality HMD because of dyskinesia of the head and neck, resulting in dizziness when viewing his videos. He used the HMD for 1 week and completed the rest of the intervention using a flat-screen device. No other adverse events occurred during the intervention or assessment.

Figure 2. Flowchart of participants through the study.**Table 2.** Adherence to the intervention.

	Video viewings (% completed)	Physical practice (% completed)
P1	105	107
P2	143	143
P3 ^a	18	23
P4	53	50
P5	105	113
P6	153	187
P7	73	97
P8 ^b	117	157
P9 ^c	65	107
P10 ^c	8	17
Value, mean (SD)	84 (49)	100 (56)

^aThis participant withdrew because of medical reasons unrelated to the trial.

^bThis participant completed the majority of the intervention using a flat-screen device.

^cThis participant did not receive the complete intervention because of COVID-19 lockdown.

Acceptability

Results from the modified PENS Questionnaire (Tables 3 and 4) suggested that the majority of participants found the use of the virtual reality system to view their videos interesting and

enjoyable. Participants experienced minimal motion sickness and would choose to use the virtual reality system to manage their freezing of gait in the future. Results for competence, presence or immersion, and intuitive controls were overall positive but variable between participants.

Table 3. Results from the modified Players Experience of Need Satisfaction Questionnaire.

Modified Players Experience of Need Satisfaction Questionnaire question	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Interest or enjoyment										
I enjoyed doing this activity very much.	7	6	4	5	4	7	7	7	7	5
This activity was fun to do.	7	5	4	6	3	7	7	6	7	5
This was not a boring activity ^a .	2	3	6	7	3	6	7	7	7	7
This activity did hold my attention ^a .	6	2	6	7	5	7	7	6	7	6
I would describe this activity as very interesting.	6	5	5	5	4	7	7	6	7	6
While I was doing this activity, I was thinking about how much I enjoyed it.	4	5	4	5	5	4	7	5	4	2
Competence										
I felt competent at using the virtual reality system to watch my videos.	6	7	5	4	5	6	6	2	7	2
I felt very capable and effective when using the virtual reality system.	6	7	5	3	6	6	6	2	7	2
Presence or immersion										
When watching my videos, I felt transported to another time and place.	1	5	2	6	6	4	7	2	1	5
When watching my videos, I felt as if I was actually there.	5	5	5	4	6	6	7	5	7	5
Watching my videos was engaging.	6	5	4	5	5	6	7	5	4	5
I experienced feelings as deeply in my videos as I have in real life.	4	4	3	5	5	3	7	5	4	3
I experienced genuine pride when I watched my videos.	4	4	2	4	3	6	2	6	4	1
I did not experience distress when I watched my videos ^a .	7	7	6	7	6	7	1	4	7	7
Intuitive controls										
Learning to use the virtual reality system was easy.	3	7	5	4	5	3	6	3	7	2
The virtual reality system controls were intuitive.	3	7	4	4	5	4	6	3	4	2
Future participation										
If given the opportunity, I would use virtual reality systems for the management of my freezing of gait in the future.	5	7	2	5	5	4	7	6	7	7
Motion sickness										
I did not experience motion sickness while using the virtual reality system ^a .	7	7	6	7	6	7	7	5	7	7

^aThis question was rephrased so that its effect direction is similar to that of the other questions to allow for the calculation of mean and SD (Multimedia Appendix 2).

Table 4. Overall results from each category from the modified Players Experience of Need Satisfaction Questionnaire^a.

Category	Value, mean (SD)
Interest or enjoyment	5.6 (1.5)
Competence	5.0 (1.8)
Presence or immersion	4.7 (1.7)
Intuitive controls	4.4 (1.6)
Future participation	5.5 (1.6)
Motion sickness	6.6 (0.7)

^aOverall mean: 5.3 (SD 1.5).

All the participants completed semistructured interviews after the intervention. Each interview lasted approximately 45-60 minutes.

Five themes emerged from the interviews. An overview of these results is presented below, whereas codes and further supporting quotes for each theme are presented in [Multimedia Appendix 5](#).

1. Reflections when seeing myself: while most participants were embarrassed or shocked by their appearances in the videos, these feelings motivated them to change. These participants used the videos to identify areas of improvement to correct their posture and gait. A total of 1 participant, who wanted to hide his disease from others, was disappointed, as he felt that he was not achieving this as well as he thought. A total of 2 participants were unconcerned with their appearances and solely focused on using the videos to learn the movement strategies to help them overcome their freezing of gait:

One of the things that I changed, looking at the video, my posture and my walking, struck me as pretty poor. I have to make conscious decision, effort, to improve. [P5]

2. My experience of using the virtual reality system: a total of 5 participants needed assistance from their carers to don or doff the HMD and navigate the platform to view their videos. Participants felt frustrated if they could not easily access their videos or when they made mistakes when using the controller to make selections. Although their abilities to use the system improved with practice, frustration when using the HMD decreased participants' motivation to engage with the intervention:

It strikes me that the virtual reality system is not foolproof; if you make a mistake in one menu or something, you have to find your way back. And any frustration like that, you don't want to carry on. [P10]

On the other hand, some participants found using the system easy and straightforward. Pre-existing technological literacy in participants might account for individual experiences when using the virtual reality system. Participants who reported feeling comfortable using technological devices found it easy to use the virtual reality system. In contrast, participants who were unfamiliar and not confident with technology reported difficulties.

Although participants were provided with personalized instruction and a printed user guide, most participants requested further resources and support to use virtual reality, such as video tutorials.

3. The role of the virtual reality system in supporting my learning: most participants found the use of the virtual reality system to be beneficial. The virtual reality system was seen as a novel and engaging tool to learn movement strategies. Using the HMD minimized distractions in the environment and allowed participants to concentrate on viewing the videos. However, a total of 2 participants did not perceive the virtual reality system as superior to flat-screen devices such as tablets, stating cost and usability as barriers to using the system:

There're less outside influences to bother you. Because if I'm looking at a computer screen...I mean there're other things on the desk, on the table, in the background. Whereas this headset excludes all that out of the equation. [P9]

I don't think it's any better than say someone holding a phone or a video and watching, so if there is a cost to it which obviously there is quite a considerable amount of money involved. [P3]

4. Developing a deeper understanding on how to manage my freezing of gait: all participants found that viewing their videos was helpful. Most participants used their videos as a guide to learn how to overcome their freezing of gait, whereas 1 participant used his video to work out how to hide his Parkinson disease. Several participants valued viewing themselves in the videos as it best reflected their individual disease presentation and personalized strategies to overcome their freezing. Participants also saw videos as exemplars for physical practice:

You got that the video of what you're supposed to be doing as you are doing it. You tell me what to do, and I didn't have the video, I don't know—I might think I'm doing it the right way. And you come back in a week's time and say that's not what I meant at all. So I'm wrong—if I have the video it shows me what I'm meant to do. [P2]

The person you can most identify with. And that's yourself. The most appropriate person is the person who's got Parkinson's themselves and is functioning. And you can see what the program is doing for you, and with you. [P9]

Most participants found the video viewing and physical practice repetitive but worthwhile. Overall, repetitive practice was acceptable as a way of reinforcing good performance, and participants reported feeling better equipped to overcome their freezing of gait when it was triggered. However, repetitive physical practice appeared to be more accepted by participants than repetitive video viewing. Several participants suggested reducing the amount of video viewing, especially once they had mastered the movement strategies. None of the participants reported issues related to the amount of physical practice:

I think it's probably the repetitiveness that is helping you without you being aware. The fact that you have done it so many times makes it easier. [P6]

I think it prepares more the attitude to when (freezing of gait) happens. For instance, you don't cut it completely or diminish. But when it happens, you get out of the situation much quicker. [P5]

5. Impact of the intervention on my daily activities: several participants reported improvements in their ability to manage their freezing of gait in other tasks and contexts. In total, 1 participant reported an improved ability to manage his freezing of gait outside of his home. He was able to board and alight from a ferry after having previously avoided public transportation. A total of 2 participants reported an improved ability to manage their freezing of gait when they were off their Parkinson disease medication. Participants reported being more

confident and less anxious when they performed tasks that typically triggered their freezing of gait:

I've got more confidence. I don't panic as bad as I used to. I'm not afraid to walk around the house or outside. This is a really great help. [P8]

In the freezing, yes. I experienced also, improvement. I got, the times that I got more intense, or the freezing when I wake up at night and I'm not with the medicine obviously. At that moment, almost every time is freezing but I start I stomp on the floor a little bit to start walking and so I applied the technique. [P5]

You know, this is a tool. By using it and doing things, (P7) can improve the way he does things around the house. The amount of falls he's had, I think have been reduced. His confidence has been a lot better. [P7 carer]

Secondary Outcome Measures

Overview

A total of 6 participants achieved their freezing of gait goal or reported performing somewhat better than their goal after receiving the intervention. A total of 3 participants reported no change. Of these, 2 participants did not receive complete intervention because of the COVID-19 lockdown (Table 5).

Table 5. Goal Attainment Scale.

ID	Goal	Performed worse than baseline (-2)	Performed at baseline level of performance (-1)	Achieved goal as expected (0)	Performed somewhat better than expected (1)	Performed much better than expected (2)
P1	Turning around to sit on a chair, toilet, or car			✓		
P2	Walking out of the unit through the front door				✓	
P3 ^a	Turning and stepping backwards at night					
P4	Turning around to sit down on the toilet (tight space)				✓	
P5	Taking the first step after standing up			✓		
P6	Walking from the bathroom to the kitchen when I am OFF my medication			✓		
P7	Walking in and out of the bathroom and laundry (tight space)				✓	
P8	Turning when opening or closing the fridge		✓			
P9 ^b	Experiencing freezing of gait in the study		✓			
P10 ^b	Experiencing freezing of gait in the toilet		✓			

^aThis participant withdrew because of medical reasons unrelated to the trial.

^bThis participant did not receive the complete intervention because of COVID-19 lockdown.

Overall, there were minimal changes to the freezing of gait, mobility, or anxiety measures from baseline to postintervention. However, there was substantial variability among participants (Multimedia Appendix 6). As this pilot study was not sufficiently powered to detect pre-post group differences, the results of 2 participants who were similar at baseline but demonstrated different levels of engagement with the intervention are highlighted below in more detail.

Participant 6

P6 was a 75-year-old female with mild to moderate disease. She had good cognition and mobility but significant freezing of gait, which was predominantly during her *off* phase. P6 also

reported high levels of anxiety. She was highly engaged and responded well to the intervention, with multiple positive outcomes.

P6 viewed and practiced 2 different tasks. In her first video, she walked from her bathroom to her kitchen. This involved making a number of turns (including a 180° turn), navigating 2 doorways, and walking along a narrow passageway. The strategy she implemented was stepping in time to a metronome at 80 beats per minute (70% of her usual cadence measured in her baseline 10-meter walk test). This strategy was selected because she reported enjoying dancing and music and had previously responded well to rhythmic music. In her second video, she practiced a more complex task of making a cup of tea in her

kitchen. This involved making a number of 180° turns while performing dual tasks, that is, holding a cup. The strategy that she implemented was counting her steps (*one, two, one, two, ...*), especially during turning. The metronome was not used in this task because of multiple noncontinuous segments. Dual tasks were also simplified into single tasks, where P6 would move the cup of tea along the bench top and take sideway steps separately so that she did not have to walk and hold onto a hot beverage concurrently (Multimedia Appendix 7).

Overall, P6 had a deep understanding of the rationale for the intervention. During her interview, she reported:

Because I visualized it, I watched it first, and it was in my head, it made it easier when I got up to move along doing the same actions. In that way it helped when I actually had to do it. I just sort of didn't have to think. Because it was like in my head. And so I just moved, you know, the way I actually visualized it.

P6 also recognized the importance of repetitive practice. She demonstrated high levels of adherence to the intervention and completed more practice than what was prescribed (video viewing=153%; physical practice=187%). Although she did not consider the virtual reality system easy or intuitive to use (PENS intuitive control=3.5/7), she felt competent using the system (PENS competence=6.0/7) and reported using the system to view her videos as interesting and enjoyable (PENS interest or enjoyment=6.3/7).

Her goal was to reduce the severity of her freezing of gait when walking from her bedroom to her kitchen before her first dose of Parkinson disease medication every morning, and she reported achieving this goal after the intervention. At postintervention, she had a reduction in percent time frozen (from 94% to 84%) and time taken to complete the Ziegler test (from 489 s to 393 s) during *off* testing. She also had a significant reduction in her anxiety score (PAS from 18/48 to 3/48), and her freezing of gait was triggered less often (CFOG-Q from 31/48 to 17/48).

Participant 4

P4 was a 78-year-old male with mild to moderate disease. He had good mobility but reported significant freezing of gait, which was also predominantly during his *off* phase. Compared with P6, he had poorer executive function (based on the Trail Making Test part B) [53] and had lower levels of anxiety. He appeared less engaged and did not respond well to the intervention.

P4 viewed and practiced 3 similar but progressively more difficult tasks. In his first video, he practiced getting on and off the toilet, which involved walking and turning in a narrow space. In his second video, he practiced getting on and off a low, soft, and deep couch. In his third video, he practiced standing up and walking away from a chair at his desk. In all 3 videos, he used movement strategies of counting his steps (*one, two, one, two, ...*) and marching on the spot when turning. P4 had not tried any movement strategies before this study, and these strategies were selected as he found them easy to implement.

Although he recognized that the movement strategies were helpful and valued seeing himself in his videos so that he could

improve his posture and walking, P4 had a low level of engagement with the intervention overall. He did not feel competent using the virtual reality system (PENS competence=3.5/7) and reported experiencing difficulties using the virtual reality system in his interview:

Because I had these problems, I sort of I don't know how to fix it so I'll wait till (the physiotherapist) comes or ring (the physiotherapist).

He also demonstrated low levels of adherence to the intervention (video viewing=53%; physical practice=50%). Interestingly, he stated several times he wished there were more applications on the virtual reality system he could explore:

I would have liked to, I think you can play games on it, can't you?

His goal was to reduce the severity of his freezing of gait when turning to sit down on the toilet, and he reported achieving this goal somewhat better than expected after the intervention. His perception of his performance contrasted with his postintervention outcome measures, as there were small increases in percent time frozen (from 22% to 27%) and time taken to complete the Ziegler test (from 37 s to 51 s) during *off* testing. His other freezing of gait measures and his anxiety also increased (NFOG-Q from 15/28 to 27/28, CFOG-Q from 12/48 to 21/48, and PAS from 10/48 to 19/48).

Discussion

Principal Findings

To the best of our knowledge, this pilot study is the first of its kind to examine the feasibility and acceptability of video self-modeling (delivered via a virtual reality HMD) plus physical practice to help people with Parkinson disease manage their freezing of gait. The results of this study showed that the intervention is a feasible and acceptable option for addressing freezing of gait in people with Parkinson disease.

This intervention was safe to deliver with no participants experiencing falls during the intervention or assessments, despite freezing of gait, meaning that participants were at a high risk of falls [37]. The physiotherapist who delivered the intervention supervised the physical practice during the home visits until the participant was deemed safe to practice either independently or with a trained carer. Future larger randomized studies of video self-modeling plus physical practice to address freezing of gait should consider investigating falls as an outcome given the potential for this intervention to ameliorate fall risk.

As virtual reality motion sickness (due to sensory conflicts between visual, vestibular, and proprioceptive inputs) was a possible adverse event, susceptibility to motion sickness was an exclusion criterion. Although all participants were screened for motion sickness, 1 participant gradually developed intolerance of the HMD after 1 week. This was because of an exacerbation of his head and neck dyskinesia when wearing the HMD, which contributed to motion sickness. The camera and virtual reality system used in this study were commercially available devices selected for their accessibility to clinicians and relatively low cost. Although virtual reality motion sickness

may be minimized by improving the technical aspects of the viewing experience, such as 6 degrees of freedom tracking, using dynamic depth of field, and providing multimodal sensory information [54], these improvements would require further development and involve higher costs. Future research to determine the incidence and intensity of virtual reality motion sickness, as well as the impact of varying technical aspects and costs of virtual reality systems, is needed in this population. Nevertheless, the use of a flat-screen device was found to be acceptable to the participant who required it.

The retention rate and adherence to intervention were high in this study. Participants completed most of the prescribed video self-modeling and physical practice sessions, with adherence rates comparable with those of other home-based exercise interventions [55]. Interestingly, the adherence rate of video viewing was less than that of physical practice. It was likely that repetitive viewings of the videos were more tedious compared with physical practice and that participants placed more value on physical practice over video viewing, even though both aspects of this intervention were designed to be complementary. This interpretation is supported by interview data, where several participants suggested reducing the amount of video viewing but did not suggest reducing the amount of physical practice. The amount and distribution of video viewing and physical practice should be further investigated in future studies to determine the best combination to improve performance while remaining engaging.

When comparing the recruitment rate of this study with others specifically targeting freezing of gait [28,56-59], our recruitment rate was lower. This was likely because of our recruitment from a general database of people with Parkinson disease, where a large proportion of individuals did not meet the inclusion criteria for this pilot study. It is also possible that potential participants with freezing of gait were deterred by the time needed to commit to the study and unfamiliarity with the technology.

This intervention was widely accepted. Most of the participants considered the intervention appropriate for their needs and would choose to do this again in the future. Although video self-modeling may be a useful tool for learning movement strategies, its use in people with Parkinson disease should be carefully considered. When people with Parkinson disease view themselves, they can bring up powerful emotions that are both positive and negative. Video self-modeling may improve self-efficacy by providing an opportunity for individuals to see their potential to overcome their freezing of gait. On the other hand, other forms of observational learning may need to be considered for people with Parkinson disease who are uncomfortable viewing themselves. Future studies are required to determine whether video self-modeling plus physical practice is superior to more generic observational learning plus physical practice for addressing freezing of gait.

Difficulties operating and navigating the virtual reality system to view videos may hinder the learning of the movement strategies and reduce adherence to the intervention. People with Parkinson disease may have tremors and dyskinesia, which makes it difficult to use a small controller, especially when vision is eliminated. They may also have difficulties donning

and doffing an HMD that requires multiple adjustments to fit securely. In addition, people with significant cognitive impairments are likely to have issues using and engaging with a complex virtual reality system. Simple and intuitive platforms that provide people with Parkinson disease with large icons for selection and sufficient time to respond facilitates navigation of the applications.

Overall, the interpretation of secondary outcomes was limited by the small sample size and the variability of the results. Although there were minimal changes in the severity of freezing of gait, this intervention may be effective in reducing anxiety in participants with high levels of anxiety. Of the 5 participants with significant levels of anxiety at baseline (PAS>14/48), 4 participants (P6, P7, P8, and P10) had significant reductions in their anxiety scores. One additional participant (P2), who did not meet the cut-off for significant anxiety, also demonstrated a significant reduction in his anxiety score (PAS from 9/48 to 0/48). This finding is supported by interview data, in which participants reported increased confidence in performing tasks where they previously experienced freezing of gait. They also reported feeling less anxious when they experienced freezing of gait and felt better equipped with strategies to overcome freezing when it was triggered.

Given that anxiety is a risk factor and a significant predictor of developing freezing of gait [18], reducing anxiety may alleviate freezing of gait by reducing the load on attentional or cognitive resources needed to control gait in people with Parkinson disease. Our results suggest that individuals with high levels of anxiety regarding their freezing of gait may be more suited and more likely to benefit from this intervention. It also suggests that personalized rehabilitation for gait impairments is required, where individual characteristics help determine intervention modality and delivery to achieve the best outcomes [60,61].

Interestingly, 2 participants in this study (P5 and P6) specifically identified tasks and set goals to overcome their freezing in their *off* phase, as they experienced minimal freezing of gait during their *on* phase when their medication was working optimally. Even though the intervention was completed when the participants were *on*, the effects of the intervention appeared to carry over to the *off* phase. In their interviews, both participants reported that they were able to successfully implement their personalized strategies when they experienced freezing and reported achieving their goals on the Goal Attainment Scale.

Given its episodic nature and the myriad of factors that can trigger or alleviate freezing of gait, the assessment of freezing is exceptionally challenging. Common outcome measures used to assess the severity of freezing of gait, such as the Freezing of Gait Questionnaire [62] and NFOG-Q, have limitations because of their subjective nature and poor responsiveness and should be used in conjunction with other freezing of gait measures [24,31]. Several objective clinical measures of freezing of gait have been developed to complement self-report measures but freezing of gait is still not reliably captured, especially in the home environment [32]. The freezing of gait measures used in this study provided a good indication of the participants' severity of freezing but might not be sufficiently sensitive to reflect the frequency and severity of freezing in daily life. Given

that this intervention was delivered in participants' homes with a highly personalized approach, the development of protocols and technology to measure freezing throughout the day in the home environment using wearable sensors is urgently required.

Conclusions

Video self-modeling using an immersive virtual reality HMD plus physical practice of personalized movement strategies is

a feasible and acceptable method of addressing freezing of gait in people with Parkinson disease. Future larger randomized controlled trials could explore the use of wearable sensors to measure freezing of gait at home for home-based intervention, the use of different platforms to deliver video self-modeling, and the impact of this intervention on freezing of gait, anxiety, and falls.

Acknowledgments

The authors would like to thank their colleagues from KU Leuven Pieter Ginis, Alice Nieuwboer, Moran Gilat, Nicholas D'cruz, Clara De Somer, and Demi Zoetewei for sharing information and useful discussions while developing video tagging protocols. They also thank Edward Gorgon for data collection and the participants for their time.

This study was supported by the University of Sydney Charles Perkins Centre Active Aging Seeding Grant. The funding body had no role in the design and conduct of the study, analysis and interpretation of the data, or review and approval of the manuscript.

Authors' Contributions

LG was involved in research conception and execution, data analysis, writing of the first draft, and manuscript review. NEA and CGC were involved in research conception and execution, data analysis, and manuscript review. NA was involved in research conception, data analysis, and manuscript review. KAEM, JS, and LC were involved in data analysis and manuscript review. SJGL and HGM were involved in manuscript review.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Intervention protocol.

[[DOCX File , 24 KB - formative_v5i11e28315_app1.docx](#)]

Multimedia Appendix 2

Modified Players Experience of Need Satisfaction Questionnaire.

[[DOCX File , 20 KB - formative_v5i11e28315_app2.docx](#)]

Multimedia Appendix 3

Semistructured interview guide.

[[DOCX File , 18 KB - formative_v5i11e28315_app3.docx](#)]

Multimedia Appendix 4

Protocol for determining percent time frozen.

[[DOCX File , 27 KB - formative_v5i11e28315_app4.docx](#)]

Multimedia Appendix 5

Interview themes, codes, and quotes.

[[DOCX File , 28 KB - formative_v5i11e28315_app5.docx](#)]

Multimedia Appendix 6

Table of individual secondary outcome data.

[[PDF File \(Adobe PDF File\), 112 KB - formative_v5i11e28315_app6.pdf](#)]

Multimedia Appendix 7

Example of self-modeling video.

[[MP4 File \(MP4 Video\), 23183 KB - formative_v5i11e28315_app7.mp4](#)]

References

1. Dorsey ER, Elbaz A, Nichols E, Abd-Allah F, Abdelalim A, Adsuar JC, et al. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol* 2018 Nov 09;17(11):939-953. [doi: [10.1016/S1474-4422\(18\)30295-3](https://doi.org/10.1016/S1474-4422(18)30295-3)] [Medline: [30287051](https://pubmed.ncbi.nlm.nih.gov/30287051/)]
2. Jankovic J. Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry* 2008 Apr 01;79(4):368-376. [doi: [10.1136/jnnp.2007.131045](https://doi.org/10.1136/jnnp.2007.131045)] [Medline: [18344392](https://pubmed.ncbi.nlm.nih.gov/18344392/)]
3. Giladi N, Nieuwboer A. Understanding and treating freezing of gait in Parkinsonism, proposed working definition, and setting the stage. *Mov Disord* 2008 Jul 30;23 Suppl 2(S2):423-425. [doi: [10.1002/mds.21927](https://doi.org/10.1002/mds.21927)] [Medline: [18668629](https://pubmed.ncbi.nlm.nih.gov/18668629/)]
4. Nutt JG, Bloem BR, Giladi N, Hallett M, Horak FB, Nieuwboer A. Freezing of gait: moving forward on a mysterious clinical phenomenon. *Lancet Neurol* 2011 Aug;10(8):734-744. [doi: [10.1016/s1474-4422\(11\)70143-0](https://doi.org/10.1016/s1474-4422(11)70143-0)]
5. Allen N, Schwarzel K, Canning G. Recurrent falls in Parkinson's disease: a systematic review. *Parkinsons Dis* 2013;2013:906274 [FREE Full text] [doi: [10.1155/2013/906274](https://doi.org/10.1155/2013/906274)] [Medline: [23533953](https://pubmed.ncbi.nlm.nih.gov/23533953/)]
6. Walton CC, Shine JM, Hall JM, O'Callaghan C, Mowszowski L, Gilat M, et al. The major impact of freezing of gait on quality of life in Parkinson's disease. *J Neurol* 2015 Jan 16;262(1):108-115. [doi: [10.1007/s00415-014-7524-3](https://doi.org/10.1007/s00415-014-7524-3)] [Medline: [25319020](https://pubmed.ncbi.nlm.nih.gov/25319020/)]
7. Dorsey ER, Sherer T, Okun MS, Bloem BR. The emerging evidence of the Parkinson pandemic. *J Parkinsons Dis* 2018 Dec 18;8(s1):3-8. [doi: [10.3233/jpd-181474](https://doi.org/10.3233/jpd-181474)]
8. Perez-Lloret S, Negre-Page L, Damier P, Delval A, Derkinderen P, Destée A, et al. Prevalence, determinants, and effect on quality of life of freezing of gait in Parkinson disease. *JAMA Neurol* 2014 Jul 01;71(7):884-890. [doi: [10.1001/jamaneurol.2014.753](https://doi.org/10.1001/jamaneurol.2014.753)] [Medline: [24839938](https://pubmed.ncbi.nlm.nih.gov/24839938/)]
9. Martens KA, Hall JM, Georgiades MJ, Gilat M, Walton CC, Matar E, et al. The functional network signature of heterogeneity in freezing of gait. *Brain* 2018 Apr 01;141(4):1145-1160. [doi: [10.1093/brain/awy019](https://doi.org/10.1093/brain/awy019)] [Medline: [29444207](https://pubmed.ncbi.nlm.nih.gov/29444207/)]
10. Cucca A, Biagioni MC, Fleisher JE, Agarwal S, Son A, Kumar P, et al. Freezing of gait in Parkinson's disease: from pathophysiology to emerging therapies. *Neurodegener Dis Manag* 2016 Oct;6(5):431-446. [doi: [10.2217/nmt-2016-0018](https://doi.org/10.2217/nmt-2016-0018)] [Medline: [27599588](https://pubmed.ncbi.nlm.nih.gov/27599588/)]
11. Lewis SJ, Shine JM. The next step: a common neural mechanism for freezing of gait. *Neuroscientist* 2016 Feb 14;22(1):72-82. [doi: [10.1177/1073858414559101](https://doi.org/10.1177/1073858414559101)] [Medline: [25398230](https://pubmed.ncbi.nlm.nih.gov/25398230/)]
12. Lewis SJ, Barker RA. A pathophysiological model of freezing of gait in Parkinson's disease. *Parkinsonism Relat Disord* 2009 Jun;15(5):333-338. [doi: [10.1016/j.parkreldis.2008.08.006](https://doi.org/10.1016/j.parkreldis.2008.08.006)] [Medline: [18930430](https://pubmed.ncbi.nlm.nih.gov/18930430/)]
13. Weiss D, Schoellmann A, Fox MD, Bohnen NI, Factor SA, Nieuwboer A, et al. Freezing of gait: understanding the complexity of an enigmatic phenomenon. *Brain* 2020 Jan 01;143(1):14-30 [FREE Full text] [doi: [10.1093/brain/awz314](https://doi.org/10.1093/brain/awz314)] [Medline: [31647540](https://pubmed.ncbi.nlm.nih.gov/31647540/)]
14. Lichter DG, Benedict RH, Hershey LA. Freezing of gait in Parkinson's disease: risk factors, their interactions, and associated nonmotor symptoms. *Parkinsons Dis* 2021 Jan 12;2021:8857204 [FREE Full text] [doi: [10.1155/2021/8857204](https://doi.org/10.1155/2021/8857204)] [Medline: [33505652](https://pubmed.ncbi.nlm.nih.gov/33505652/)]
15. Heremans E, Nieuwboer A, Spildooren J, Vandenbossche J, Deroost N, Soetens E, et al. Cognitive aspects of freezing of gait in Parkinson's disease: a challenge for rehabilitation. *J Neural Transm* 2013 Apr 18;120(4):543-557. [doi: [10.1007/s00702-012-0964-y](https://doi.org/10.1007/s00702-012-0964-y)] [Medline: [23328947](https://pubmed.ncbi.nlm.nih.gov/23328947/)]
16. Pimenta M, Moreira D, Nogueira T, Silva C, Pinto EB, Valenca GT, et al. Anxiety independently contributes to severity of freezing of gait in people with Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 2019 Jan;31(1):80-85. [doi: [10.1176/appi.neuropsych.17090177](https://doi.org/10.1176/appi.neuropsych.17090177)] [Medline: [30187821](https://pubmed.ncbi.nlm.nih.gov/30187821/)]
17. Witt I, Ganjavi H, MacDonald P. Relationship between freezing of gait and anxiety in Parkinson's disease patients: a systemic literature review. *Parkinsons Dis* 2019 Jul 24;2019:6836082 [FREE Full text] [doi: [10.1155/2019/6836082](https://doi.org/10.1155/2019/6836082)] [Medline: [31428304](https://pubmed.ncbi.nlm.nih.gov/31428304/)]
18. Martens KA, Peterson DS, Almeida QJ, Lewis SJ, Hausdorff JM, Nieuwboer A. Behavioural manifestations and associated non-motor features of freezing of gait: a narrative review and theoretical framework. *Neurosci Biobehav Rev* 2020 Sep;116:350-364. [doi: [10.1016/j.neubiorev.2020.06.026](https://doi.org/10.1016/j.neubiorev.2020.06.026)] [Medline: [32603716](https://pubmed.ncbi.nlm.nih.gov/32603716/)]
19. Tomlinson C, Patel S, Meek C, Herd CP, Clarke CE, Stowe R, et al. Physiotherapy versus placebo or no intervention in Parkinson's disease. *Cochrane Database Syst Rev* 2013 Sep 10(9):CD002817 [FREE Full text] [doi: [10.1002/14651858.CD002817.pub4](https://doi.org/10.1002/14651858.CD002817.pub4)] [Medline: [24018704](https://pubmed.ncbi.nlm.nih.gov/24018704/)]
20. Delgado-Alvarado M, Marano M, Santurtún A, Urriaga-Gallano A, Tordesillas-Gutierrez D, Infante J. Nonpharmacological, nonsurgical treatments for freezing of gait in Parkinson's disease: a systematic review. *Mov Disord* 2020 Feb 26;35(2):204-214. [doi: [10.1002/mds.27913](https://doi.org/10.1002/mds.27913)] [Medline: [31769904](https://pubmed.ncbi.nlm.nih.gov/31769904/)]
21. Cosentino C, Baccini M, Putzolu M, Ristori D, Avanzino L, Pelosin E. Effectiveness of physiotherapy on freezing of gait in Parkinson's disease: a systematic review and meta-analyses. *Mov Disord* 2020 Apr 04;35(4):523-536. [doi: [10.1002/mds.27936](https://doi.org/10.1002/mds.27936)] [Medline: [31799734](https://pubmed.ncbi.nlm.nih.gov/31799734/)]
22. Rutz DG, Benninger DH. Physical therapy for freezing of gait and gait impairments in Parkinson disease: a systematic review. *PM R* 2020 Nov 19;12(11):1140-1156. [doi: [10.1002/pmrj.12337](https://doi.org/10.1002/pmrj.12337)] [Medline: [31994842](https://pubmed.ncbi.nlm.nih.gov/31994842/)]

23. Ginis P, Nackaerts E, Nieuwboer A, Heremans E. Cueing for people with Parkinson's disease with freezing of gait: a narrative review of the state-of-the-art and novel perspectives. *Ann Phys Rehabil Med* 2018 Nov;61(6):407-413 [FREE Full text] [doi: [10.1016/j.rehab.2017.08.002](https://doi.org/10.1016/j.rehab.2017.08.002)] [Medline: [28890341](https://pubmed.ncbi.nlm.nih.gov/28890341/)]
24. Hulzinga F, Nieuwboer A, Dijkstra BW, Mancini M, Strouwen C, Bloem BR, et al. The New Freezing of Gait Questionnaire: unsuitable as an outcome in clinical trials? *Mov Disord Clin Pract* 2020 Feb 14;7(2):199-205 [FREE Full text] [doi: [10.1002/mdc3.12893](https://doi.org/10.1002/mdc3.12893)] [Medline: [32071940](https://pubmed.ncbi.nlm.nih.gov/32071940/)]
25. Pelosin E, Avanzino L, Bove M, Stramesi P, Nieuwboer A, Abbruzzese G. Action observation improves freezing of gait in patients with Parkinson's disease. *Neurorehabil Neural Repair* 2010 Oct 07;24(8):746-752. [doi: [10.1177/1545968310368685](https://doi.org/10.1177/1545968310368685)] [Medline: [20453155](https://pubmed.ncbi.nlm.nih.gov/20453155/)]
26. Pelosin E, Barella R, Bet C, Magioncalda E, Putzolu M, Di Biasio F, et al. Effect of group-based rehabilitation combining action observation with physiotherapy on freezing of gait in Parkinson's disease. *Neural Plast* 2018 May 27;2018:4897276 [FREE Full text] [doi: [10.1155/2018/4897276](https://doi.org/10.1155/2018/4897276)] [Medline: [29977280](https://pubmed.ncbi.nlm.nih.gov/29977280/)]
27. Agosta F, Gatti R, Sarasso E, Volonté MA, Canu E, Meani A, et al. Brain plasticity in Parkinson's disease with freezing of gait induced by action observation training. *J Neurol* 2017 Jan 24;264(1):88-101. [doi: [10.1007/s00415-016-8309-7](https://doi.org/10.1007/s00415-016-8309-7)] [Medline: [27778161](https://pubmed.ncbi.nlm.nih.gov/27778161/)]
28. Mezzarobba S, Grassi M, Pellegrini L, Catalan M, Kruger B, Furlanis G, et al. Action observation plus sonification. A novel therapeutic protocol for Parkinson's patient with freezing of gait. *Front Neurol* 2017 Jan 4;8:723 [FREE Full text] [doi: [10.3389/fneur.2017.00723](https://doi.org/10.3389/fneur.2017.00723)] [Medline: [29354092](https://pubmed.ncbi.nlm.nih.gov/29354092/)]
29. Abbruzzese G, Avanzino L, Marchese R, Pelosin E. Action observation and motor imagery: innovative cognitive tools in the rehabilitation of Parkinson's disease. *Parkinsons Dis* 2015;2015:124214 [FREE Full text] [doi: [10.1155/2015/124214](https://doi.org/10.1155/2015/124214)] [Medline: [26495150](https://pubmed.ncbi.nlm.nih.gov/26495150/)]
30. Nieuwboer A, Rochester L, Herman T, Vandenberghe W, Emil GE, Thomaes T, et al. Reliability of the New Freezing of Gait Questionnaire: agreement between patients with Parkinson's disease and their carers. *Gait Posture* 2009 Nov;30(4):459-463. [doi: [10.1016/j.gaitpost.2009.07.108](https://doi.org/10.1016/j.gaitpost.2009.07.108)] [Medline: [19660949](https://pubmed.ncbi.nlm.nih.gov/19660949/)]
31. Shine J, Moore S, Bolitho S, Morris T, Dilda V, Naismith S, et al. Assessing the utility of Freezing of Gait Questionnaires in Parkinson's disease. *Parkinsonism Relat Disord* 2012 Jan;18(1):25-29. [doi: [10.1016/j.parkrel.2011.08.002](https://doi.org/10.1016/j.parkrel.2011.08.002)] [Medline: [21872523](https://pubmed.ncbi.nlm.nih.gov/21872523/)]
32. Mancini M, Bloem BR, Horak FB, Lewis SJ, Nieuwboer A, Nonnekes J. Clinical and methodological challenges for assessing freezing of gait: future perspectives. *Mov Disord* 2019 Jun 02;34(6):783-790 [FREE Full text] [doi: [10.1002/mds.27709](https://doi.org/10.1002/mds.27709)] [Medline: [31046191](https://pubmed.ncbi.nlm.nih.gov/31046191/)]
33. Buggey T, Ogle L. Video self-modeling. *Psychol Schs* 2011 Dec 19;49(1):52-70. [doi: [10.1002/pits.20618](https://doi.org/10.1002/pits.20618)]
34. Dowrick PW. A review of self modeling and related interventions. *Appl Prevent Psychol* 1999 Dec;8(1):23-39. [doi: [10.1016/s0962-1849\(99\)80009-2](https://doi.org/10.1016/s0962-1849(99)80009-2)]
35. Molloy DW, Alemayehu E, Roberts R. Reliability of a standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry* 1991 Jan;148(1):102-105. [doi: [10.1176/ajp.148.1.102](https://doi.org/10.1176/ajp.148.1.102)] [Medline: [1984692](https://pubmed.ncbi.nlm.nih.gov/1984692/)]
36. Martens KA, Shine JM, Walton CC, Georgiades MJ, Gilat M, Hall JM, et al. Evidence for subtypes of freezing of gait in Parkinson's disease. *Mov Disord* 2018 Jul 28;33(7):1174-1178. [doi: [10.1002/mds.27417](https://doi.org/10.1002/mds.27417)] [Medline: [30153383](https://pubmed.ncbi.nlm.nih.gov/30153383/)]
37. Paul SS, Canning CG, Sherrington C, Lord SR, Close JC, Fung VS. Three simple clinical tests to accurately predict falls in people with Parkinson's disease. *Mov Disord* 2013 May 28;28(5):655-662. [doi: [10.1002/mds.25404](https://doi.org/10.1002/mds.25404)] [Medline: [23450694](https://pubmed.ncbi.nlm.nih.gov/23450694/)]
38. Nieuwboer A, Rochester L. Cueing gait and gait-related mobility in patients with Parkinson's disease: developing a therapeutic method based on the international classification of functioning, disability, and health. *Top Geriatr Rehabil* 2008;24(2):151-165. [doi: [10.1097/01.TGR.0000318902.80066.ae](https://doi.org/10.1097/01.TGR.0000318902.80066.ae)]
39. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. *Motiv Emot* 2006 Nov 29;30(4):344-360. [doi: [10.1007/s11031-006-9051-8](https://doi.org/10.1007/s11031-006-9051-8)]
40. Turner-Stokes L. Goal attainment scaling (GAS) in rehabilitation: a practical guide. *Clin Rehabil* 2009 Apr 29;23(4):362-370. [doi: [10.1177/0269215508101742](https://doi.org/10.1177/0269215508101742)] [Medline: [19179355](https://pubmed.ncbi.nlm.nih.gov/19179355/)]
41. Ziegler K, Schroeteler F, Ceballos-Baumann AO, Fietzek UM. A new rating instrument to assess festination and freezing gait in Parkinsonian patients. *Mov Disord* 2010 Jun 15;25(8):1012-1018. [doi: [10.1002/mds.22993](https://doi.org/10.1002/mds.22993)] [Medline: [20310009](https://pubmed.ncbi.nlm.nih.gov/20310009/)]
42. Mancini M, Smulders K, Cohen RG, Horak FB, Giladi N, Nutt JG. The clinical significance of freezing while turning in Parkinson's disease. *Neuroscience* 2017 Feb 20;343:222-228 [FREE Full text] [doi: [10.1016/j.neuroscience.2016.11.045](https://doi.org/10.1016/j.neuroscience.2016.11.045)] [Medline: [27956066](https://pubmed.ncbi.nlm.nih.gov/27956066/)]
43. Gilat M. How to annotate freezing of gait from video: a standardized method using open-source software. *J Parkinsons Dis* 2019 Oct 11;9(4):821-824. [doi: [10.3233/jpd-191700](https://doi.org/10.3233/jpd-191700)] [Medline: [31524181](https://pubmed.ncbi.nlm.nih.gov/31524181/)]
44. Herman T, Dagan M, Shema-Shiratzky S, Reches T, Brozgol M, Giladi N, et al. Advantages of timing the duration of a freezing of gait-provoking test in individuals with Parkinson's disease. *J Neurol* 2020 Sep 07;267(9):2582-2588. [doi: [10.1007/s00415-020-09856-7](https://doi.org/10.1007/s00415-020-09856-7)] [Medline: [32383040](https://pubmed.ncbi.nlm.nih.gov/32383040/)]
45. Watson MJ. Refining the ten-metre walking test for use with neurologically impaired people. *Physiotherapy* 2002 Jul;88(7):386-397. [doi: [10.1016/s0031-9406\(05\)61264-3](https://doi.org/10.1016/s0031-9406(05)61264-3)]

46. Podsiadlo D, Richardson S. The Timed Up and Go: a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 Feb 27;39(2):142-148. [doi: [10.1111/j.1532-5415.1991.tb01616.x](https://doi.org/10.1111/j.1532-5415.1991.tb01616.x)] [Medline: [1991946](https://pubmed.ncbi.nlm.nih.gov/1991946/)]
47. Leentjens AF, Dujardin K, Pontone GM, Starkstein SE, Weintraub D, Martinez-Martin P. The Parkinson Anxiety Scale (PAS): development and validation of a new anxiety scale. *Mov Disord* 2014 Jul 23;29(8):1035-1043. [doi: [10.1002/mds.25919](https://doi.org/10.1002/mds.25919)] [Medline: [24862344](https://pubmed.ncbi.nlm.nih.gov/24862344/)]
48. Goetz C, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Movement Disorder Society UPDRS Revision Task Force. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008 Nov 15;23(15):2129-2170. [doi: [10.1002/mds.22340](https://doi.org/10.1002/mds.22340)] [Medline: [19025984](https://pubmed.ncbi.nlm.nih.gov/19025984/)]
49. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967 May 01;17(5):427-442. [doi: [10.1212/wnl.17.5.427](https://doi.org/10.1212/wnl.17.5.427)] [Medline: [6067254](https://pubmed.ncbi.nlm.nih.gov/6067254/)]
50. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. *Nat Protoc* 2006 Dec 21;1(5):2277-2281. [doi: [10.1038/nprot.2006.390](https://doi.org/10.1038/nprot.2006.390)] [Medline: [17406468](https://pubmed.ncbi.nlm.nih.gov/17406468/)]
51. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. *Int J Qual Methods* 2017 Oct 02;16(1):160940691773384. [doi: [10.1177/1609406917733847](https://doi.org/10.1177/1609406917733847)]
52. Martínez-Martín P, Rodríguez-Blázquez C, Alvarez M, Arakaki T, Arillo VC, Chaná P, et al. Parkinson's disease severity levels and MDS-Unified Parkinson's Disease Rating Scale. *Parkinsonism Relat Disord* 2015 Jan;21(1):50-54. [doi: [10.1016/j.parkreldis.2014.10.026](https://doi.org/10.1016/j.parkreldis.2014.10.026)] [Medline: [25466406](https://pubmed.ncbi.nlm.nih.gov/25466406/)]
53. Naismith SL, Shine JM, Lewis SJ. The specific contributions of set-shifting to freezing of gait in Parkinson's disease. *Mov Disord* 2010 Jun 15;25(8):1000-1004. [doi: [10.1002/mds.23005](https://doi.org/10.1002/mds.23005)] [Medline: [20198644](https://pubmed.ncbi.nlm.nih.gov/20198644/)]
54. Chang E, Kim HT, Yoo B. Virtual reality sickness: a review of causes and measurements. *Int J Hum Comput Interact* 2020 Jul 02;36(17):1658-1682. [doi: [10.1080/10447318.2020.1778351](https://doi.org/10.1080/10447318.2020.1778351)]
55. Flynn A, Allen NE, Dennis S, Canning CG, Preston E. Home-based prescribed exercise improves balance-related activities in people with Parkinson's disease and has benefits similar to centre-based exercise: a systematic review. *J Physiother* 2019 Oct;65(4):189-199 [FREE Full text] [doi: [10.1016/j.jphys.2019.08.003](https://doi.org/10.1016/j.jphys.2019.08.003)] [Medline: [31521554](https://pubmed.ncbi.nlm.nih.gov/31521554/)]
56. Fietzek UM, Schroeteler FE, Ziegler K, Zwosta J, Ceballos-Baumann AO. Randomized cross-over trial to investigate the efficacy of a two-week physiotherapy programme with repetitive exercises of cueing to reduce the severity of freezing of gait in patients with Parkinson's disease. *Clin Rehabil* 2014 Sep 01;28(9):902-911. [doi: [10.1177/0269215514527299](https://doi.org/10.1177/0269215514527299)] [Medline: [24691217](https://pubmed.ncbi.nlm.nih.gov/24691217/)]
57. Martin T, Weatherall M, Anderson TJ, MacAskill MR. A randomized controlled feasibility trial of a specific cueing program for falls management in persons with Parkinson disease and freezing of gait. *J Neurol Phys Ther* 2015 Jul;39(3):179-184. [doi: [10.1097/NPT.000000000000093](https://doi.org/10.1097/NPT.000000000000093)] [Medline: [26050074](https://pubmed.ncbi.nlm.nih.gov/26050074/)]
58. Walton CC, Mowszowski L, Gilat M, Hall JM, O'Callaghan C, Muller AJ, et al. Cognitive training for freezing of gait in Parkinson's disease: a randomized controlled trial. *NPJ Parkinsons Dis* 2018 May 18;4(1):15 [FREE Full text] [doi: [10.1038/s41531-018-0052-6](https://doi.org/10.1038/s41531-018-0052-6)] [Medline: [29796409](https://pubmed.ncbi.nlm.nih.gov/29796409/)]
59. King LA, Mancini M, Smulders K, Harker G, Lapidus JA, Ramsey K, et al. Cognitively challenging agility boot camp program for freezing of gait in Parkinson disease. *Neurorehabil Neural Repair* 2020 May 04;34(5):417-427 [FREE Full text] [doi: [10.1177/1545968320909331](https://doi.org/10.1177/1545968320909331)] [Medline: [32249668](https://pubmed.ncbi.nlm.nih.gov/32249668/)]
60. Nonnekes J, Nieuwboer A. Towards personalized rehabilitation for gait impairments in Parkinson's disease. *J Parkinsons Dis* 2018 Dec 18;8(s1):101-106. [doi: [10.3233/jpd-181464](https://doi.org/10.3233/jpd-181464)] [Medline: [30584154](https://pubmed.ncbi.nlm.nih.gov/30584154/)]
61. Ryden LE, Lewis SJ. Parkinson's disease in the era of personalised medicine: one size does not fit all. *Drugs Aging* 2019 Feb 17;36(2):103-113. [doi: [10.1007/s40266-018-0624-5](https://doi.org/10.1007/s40266-018-0624-5)] [Medline: [30556112](https://pubmed.ncbi.nlm.nih.gov/30556112/)]
62. Giladi N, Shabtai H, Simon E, Biran S, Tal J, Korczyn A. Construction of Freezing of Gait Questionnaire for patients with Parkinsonism. *Parkinsonism Relat Disord* 2000 Jul 01;6(3):165-170. [doi: [10.1016/s1353-8020\(99\)00062-0](https://doi.org/10.1016/s1353-8020(99)00062-0)] [Medline: [10817956](https://pubmed.ncbi.nlm.nih.gov/10817956/)]

Abbreviations

CFOG-Q: Characterizing Freezing of Gait Questionnaire

HMD: head-mounted display

NFOG-Q: New Freezing of Gait Questionnaire

PAS: Parkinson Anxiety Scale

PENS: Players Experience of Need Satisfaction

Edited by G Eysenbach; submitted 28.02.21; peer-reviewed by K Erenli; comments to author 29.03.21; revised version received 05.04.21; accepted 18.09.21; published 03.11.21.

Please cite as:

*Goh L, Allen NE, Ahmadpour N, Ehgoetz Martens KA, Song J, Clemson L, Lewis SJG, MacDougall HG, Canning CG
A Video Self-Modeling Intervention Using Virtual Reality Plus Physical Practice for Freezing of Gait in Parkinson Disease: Feasibility and Acceptability Study*

JMIR Form Res 2021;5(11):e28315

URL: <https://formative.jmir.org/2021/11/e28315>

doi: [10.2196/28315](https://doi.org/10.2196/28315)

PMID: [34730537](https://pubmed.ncbi.nlm.nih.gov/34730537/)

©Lina Goh, Natalie E Allen, Naseem Ahmadpour, Kaylena A Ehgoetz Martens, Joeun Song, Lindy Clemson, Simon J G Lewis, Hamish G MacDougall, Colleen G Canning. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.11.2021. 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 and Acceptability of a Digital Patient-Reported Outcome Tool in Routine Outpatient Diabetes Care: Mixed Methods Formative Pilot Study

Soren E Skovlund^{1,2,3}, PhD; Lise Havbæk Troelsen³, RN; Lise Møllgaard Noergaard³, RN; Anna Pietraszek³, MD; Poul Erik Jakobsen³, MD; Niels Ejksjaer³, MD

¹Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

²Department of Endocrinology, Aalborg University Hospital, Aalborg, Denmark

³Steno Diabetes Center North Denmark, Aalborg University Hospital, Aalborg, Denmark

Corresponding Author:

Soren E Skovlund, PhD

Department of Endocrinology

Aalborg University Hospital

Mølleparkvej 4

Aalborg, 9000

Denmark

Phone: 45 40228835

Email: soren@sorenskovlund.com

Abstract

Background: Improvements in the digital capabilities of health systems provide new opportunities for the integration of patient-reported outcome (PRO) solutions in routine care, which can facilitate the delivery of person-centered diabetes care. We undertook this study as part of our development of a new digital PRO diabetes questionnaire and clinical dialog support tool for use by people with diabetes and their health care professionals (HCPs) to improve person-centered diabetes care quality and outcomes.

Objective: This study evaluates the feasibility, acceptability, and perceived benefits and impacts of using a digital PRO diabetes tool, *DiaProfil*, in routine outpatient diabetes care.

Methods: Overall, 12 people with diabetes scheduled for routine medical diabetes visits at the outpatient clinic were recruited. Purposive sampling was used to optimize heterogeneity regarding age, gender, duration, type of diabetes, treatment modality, and disease severity. Participants filled out a PRO diabetes questionnaire 2 to 5 days before their visit. During the visit, HCPs used a digital PRO tool to review PRO data with the person with diabetes for collaborative care planning. Participants completed evaluation forms before and after the visit and were interviewed for 30 to 45 minutes after the visit. HCPs completed the evaluation questionnaires after each visit. All visits were audio-recorded and transcribed for analysis. Data were analyzed using quantitative, qualitative, and mixed methods analyses.

Results: People with diabetes found the PRO diabetes questionnaire to be relevant, acceptable, and feasible to complete from home. People with diabetes and HCPs found the digital PRO tool to be feasible and acceptable for use during the diabetes visit and would like to continue using it. HCPs were able to use the tool in a person-centered manner, as intended. For several people with diabetes, completion of the questionnaire facilitated positive reflection and better preparation for the visit. The use of the PRO tool primarily improved the quality of the dialog by improving the identification and focus on the issues most important to the person with diabetes. People with diabetes did not report any negative aspects of the PRO tool, whereas HCPs highlighted that it was demanding when the person with diabetes had many PRO issues that required attention within the predefined time allocated for a visit.

Conclusions: The Danish PRO diabetes questionnaire and the digital tool, *DiaProfil*, are feasible and acceptable solutions for routine diabetes visits, and this tool may generate important benefits related to advancement of person-centered care. Further research is now required to corroborate and expand these formative insights on a larger scale and in diverse health care settings. The results of this study are therefore being used to define research hypotheses and finalize real-world PRO evaluation tools for a forthcoming large-scale multisector implementation study in Denmark.

KEYWORDS

patient-reported outcomes; diabetes; person-centered care; person-centered communication; dialog; mental health; self-management; collaborative care; value-based health care; mixed-methods; mobile phone; mHealth

Introduction

Background

Successful diabetes care requires a whole-person, collaborative care approach that focuses on an individual's biological, psychological, and social health, well-being, functioning, values, preferences, and priorities [1].

Digital patient-reported outcome (PRO) solutions for clinical practice may help improve aspects of the care experience regarding person-centered chronic illness care [2] and potential outcomes of care [3]. PRO interventions may facilitate aspects of empowerment for people with diabetes through facilitation of reflective learning regarding (1) how diabetes affects one personally, (2) one's preferred role in own care, (3) prioritization of own goals and taking an active role in developing action plans, and (4) structuring a process of ongoing experimentation and self-evaluation of action plan outcomes and efforts over time [4,5]. PRO tools can additionally function as support for health care professionals (HCPs) with regard to person-centered dialog, care planning, coordination of care, treatment decisions, treatment and outcome monitoring, psychosocial screening, and shared decision-making [6].

We developed a new digital tool, *DiaProfil*, to facilitate the use of PRO data by HCPs and people with diabetes in collaborative care visits using a formative, participatory design involving people with diabetes, family members of people with diabetes, and a multidisciplinary HCP team in all design phases [3]. Participatory research and user involvement help ensure that health care interventions are fit for the intended purpose [7-11], can be seamlessly integrated into care, and can deliver optimal public health impact [12-14]. We adopted these methods in the development of both the national PRO diabetes questionnaire and the PRO digital tool, and this pilot study was a part of the formative evaluation process.

DiaProfil provides a user-friendly mobile or internet interface for people with diabetes to complete a psychometrically designed, adaptive diabetes PRO questionnaire before their scheduled diabetes visits. This questionnaire can be completed via mobile phones, tablets, or PCs and provides a one-screen interactive overview (dashboard) of the PRO results for use by HCPs with the person with diabetes during the visit. The PRO dashboard is designed to make the PRO data available in a way that facilitates effective, collaborative, and action-orientated use of the data during the care visit. The aim is that HCPs and people with diabetes use the tool together to gain a common understanding of the overall perspective of the person with diabetes regarding their overall life with diabetes, priorities, and needs.

Aims

This study aims to evaluate the feasibility and acceptability of the PRO diabetes questionnaire and the first viable version of *DiaProfil* in routine diabetes visits at an outpatient clinic and to undertake the initial exploration of the perceived benefits related to the use of PRO by people with diabetes and HCPs. The specific research questions were as follows:

1. How do people with diabetes experience the feasibility, acceptability, relevance, comprehension, and adequacy of topic coverage of the PRO questionnaire when used as intended in the context of a routine visit?
2. How do people with diabetes and HCPs experience the use of *DiaProfil* in connection with a routine diabetes visit in relation to feasibility, acceptability, perceived benefits, clinical utility, and challenges?
3. Specifically, do people with diabetes and HCPs experience the intended and hypothesized benefits of the PRO tool in improving patient participation and quality of the dialog?

As part of the formative process, this study additionally aims to facilitate the refinement of research hypotheses and finalize evaluation questionnaires for use in future real-world studies on the implementation and effectiveness of the *DiaProfil* tool in Denmark.

Methods

This study was a formative, mixed-methods, single-arm, acceptability, feasibility pilot study that evaluated an office-based digital PRO tool intervention at an outpatient diabetes clinic.

Recruitment

The eligibility criteria were age >18 years, diagnosis of type 1 or type 2 diabetes, diabetes duration of at least 1 year, and planned attendance for a routine visit at the diabetes outpatient clinic during the study period. Exclusion criteria were severe mental illness or major cognitive or language difficulties that would prevent the ability to fill out the diabetes questionnaire.

We used purposive sampling and consecutive recruitment to maximize the representation of type and duration of diabetes, age, gender, treatment regimen, and disease severity. Eligible participants were identified by the study nurses from the electronic booking system in the diabetes clinic and invited to take part in the pilot study by telephone. The study was described as a pilot test of a diabetes questionnaire designed to help improve the quality of diabetes visits. All participants signed informed consent before study enrollment.

Study Design

The study was approved by the local institutional review board and deemed out of scope for the ethical review board because

of the absence of a clinical treatment or intervention and limited risk.

The study design and intervention included the following:

1. Each participant received a secure email with a weblink for mobile or web-based access to complete informed consent, the PRO diabetes questionnaire, and an evaluation questionnaire about the PRO questionnaire. The participant was asked to complete this questionnaire 2-10 days before their scheduled visit. If it was not possible to do it at home, the person with diabetes was encouraged to contact the clinic to arrange on-site completion.
2. HCPs were able to immediately access PRO results once completed by the person with diabetes but were asked to only access results immediately before the visit to mimic routine diabetes care.
3. At the visit, HCPs used the *DiaProfil* PRO dashboard with the person with diabetes to jointly review PRO results and plan care collaboratively around the priorities of the person with diabetes. Each visit was audio-recorded and transcribed verbatim.
4. The HCPs and people with diabetes completed the evaluation questionnaires immediately after the visit about their experiences using PRO. HCPs filled out a web-based

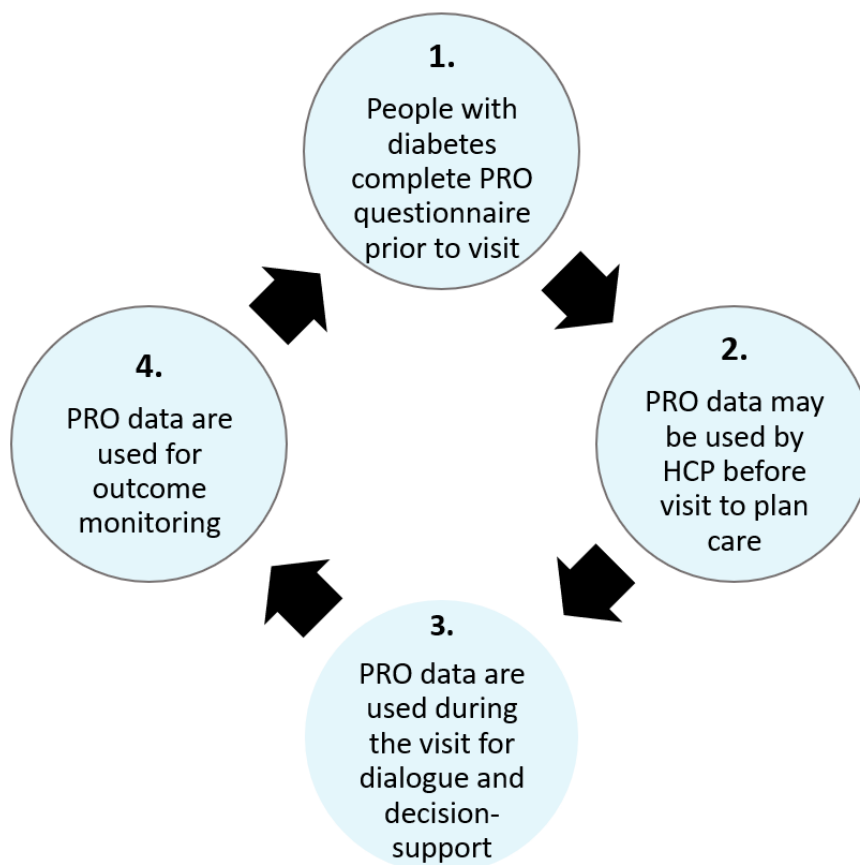
form describing any issues, errors, or concerns relating to the PRO results and how they were displayed using algorithms.

5. People with diabetes were interviewed for 30-45 minutes using a semistructured interview guide right after the diabetes visit by a researcher not involved in care for that person with diabetes. The PRO dashboard and questionnaire results were available in the interview to facilitate a detailed discussion with the person with diabetes about any feedback to the individual's PRO results, as shown on the PRO dashboard. It was emphasized upfront to people with diabetes that their feedback via questionnaires and interviews was kept strictly confidential and would not be shared with the clinical care team.

The PRO Diabetes Intervention

The aim of the PRO diabetes intervention was to increase the active participation of people with diabetes in their own care and improve the quality of the dialog between people with diabetes and HCPs, and overall care quality by focusing on optimizing value for people with diabetes [3]. The basic steps of the PRO diabetes intervention are shown in Figure 1. The digital tool *DiaProfil* was used by people with diabetes to complete the PRO questionnaire and by HCPs to manage PRO data and use the PRO data during visits.

Figure 1. Basic steps of the patient-reported outcome diabetes intervention in clinical practice. PRO: patient-reported outcome.



People with diabetes completed the PRO diabetes questionnaire by phone, tablet, or PC using *DiaProfil* at home 2-10 days before the visit. The diabetes questionnaire consisted of 33-71 items (depending on the activation of branch logic) that measured health, life situation influencing diabetes, general and

diabetes-specific social support, psychological well-being, depression, symptom distress (neuropathy pain, cardiovascular symptoms, gastrointestinal symptoms, sexual dysfunction, sleep difficulties, and foot problems), daily life with diabetes, worries about diabetes, confidence in diabetes self-management, blood

sugar regulation (including hypoglycemia and blood sugar stability), medical experience and satisfaction, access to HCPs, priority issues for support, and preferred topics to discuss. The HCPs used the PRO dashboard in *DiaProfil* with the person with diabetes during the visit to review the person’s priorities and issues and collaboratively plan care. The HCPs were recommended to review the PRO dashboard in advance, share the screen for mutual viewing, explain the PRO dashboard and the color-coding, maintain nonverbal communication and eye contact, use open-ended questions and active listening to prompt more information and confirm findings, and cover all flagged PRO issues. Our recommendations for person-centered use of the PRO tool were quite similar to recently recommended strategies for person-centered communication when using PRO data in clinical practice in other studies [15].

The Development and Design of the Digital PRO Diabetes Tool: DiaProfil

This formative study is a part of the participatory development process of the digital PRO tool *DiaProfil* by the VBHC-PRO-DIA (Value-Based Health Care and PRO in Diabetes) research team and was conducted from 2018 to 2020. We developed *DiaProfil* as a new tool to facilitate the coordinated use of the national PRO diabetes questionnaire, also developed as part of this project, in different health care settings to improve person-centered, value-based diabetes care. The PRO tool measures diabetes outcome constructs previously established as important to people with diabetes and HCP in Denmark [16].

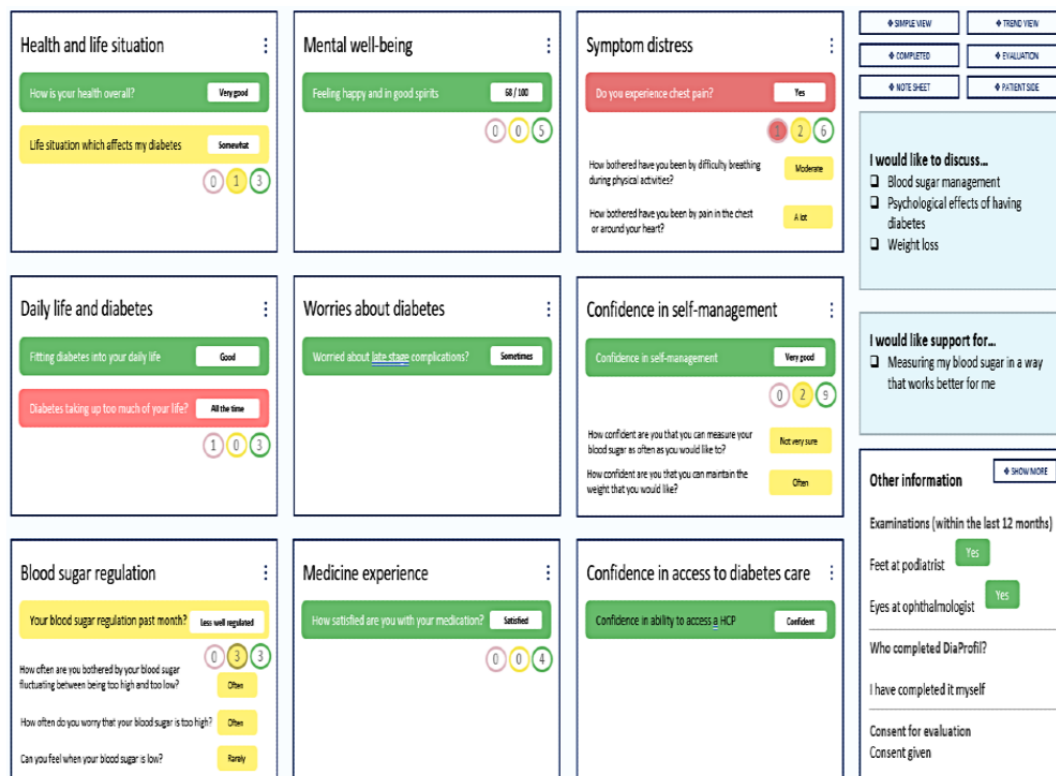
Each PRO item or scale score is shown on the *DiaProfil* HCP dashboard using a color defined by a predefined scoring algorithm. Green indicates that there is likely no problem for the person with diabetes, yellow indicates a possible issue of concern for the person with diabetes that should be considered, and red indicates that there appears to be a problem that the HCP and the person with diabetes should make sure to review and address.

The interface for people with diabetes includes a user-friendly digital interface for questionnaire completion, which was developed and tested with people with diabetes using an iterative participatory process with user-testing to optimize user-friendliness. Only one question is depicted on the screen at a time to facilitate ease of use and lower cognitive burden.

The dashboard provides a one-screen instant overview of PRO results by presenting the results in 9 main themes. On the right side of the screen, the person with diabetes’ own priorities for self-management support and topics to discuss are flagged for use as a starting point for the dialog. By clicking or touching the screen, the HCP can access dialog tips, information resources, local treatment, and referral information relevant for each PRO output.

Figure 2 shows a screenshot of the PRO dashboard with examples of how the results may be shown. Depending on individual PRO results, a variable level of information is shown on the screen.

Figure 2. Screenshot from the digital patient-reported outcome diabetes tool, *DiaProfil*. Only examples are shown. The text is abbreviated. The figure is only intended to illustrate the design of the dashboard. PRO: patient-reported outcome.



Involvement of People With Diabetes in the Research Process

A panel of people with diabetes who represented the target group for the PRO intervention were involved as partners in this study to ensure that the perspective of people with diabetes was considered at all research phases [17]. This was important to ensure relevance of the research and research questions to people with diabetes [18]. A total of 2 adults with type 2 diabetes and 3 with type 1 diabetes were part of this panel. Of these 5 participants, 3 (60%) were women, and 2 (40%) were men. The participants' age ranged from 30s to 70s, and they had different levels of disease burden. Approximately 60% (3/5) had extensive experience representing unmet needs and perspectives of people with diabetes beyond personal experiences and 40% (2/5) had some previous experience. One was also a health professional and was able to also consider the perspective of HCPs. The main method of involvement was group work meetings that involved both the user panel and members of the multidisciplinary clinical research team. This group focused on reviewing and cocreating the study aims, study design, sampling strategy, and evaluation questionnaires. In addition, this group worked with hands-on user-testing and co-designing the *DiaProfil* app tool and interface for people with diabetes. The user panel also contributed to the scientific development of parts of the PRO diabetes questionnaire.

Data Collection

Clinical charts (HbA_{1c}, cholesterol, blood pressure, and complication data), sociodemographic data, and treatment data (age, gender, duration of diabetes, type of diabetes, medical therapy, and technology use) were collected from all study

participants using chart reviews and questionnaires. PRO and evaluation questionnaires were administered to people with diabetes and HCPs using the *DiaProfil* platform. People with diabetes completed evaluation questionnaires before and after their consultation, and HCPs completed the questionnaires after the visit. Questionnaires were purposely based on qualitative data collected from workshops held with people with diabetes and HCPs as part of the development of the PRO questionnaire.

An overview of the evaluation questionnaires is shown in **Textbox 1**. Generic items were designed to evaluate the quality of autonomy-supportive and person-centered communication [19,20]. The semistructured interview guide was designed based on research questions for the study and qualitative data from involvement of people with diabetes during the intervention development process. The interview guide addressed the following main elements: (1) motivation to participate; (2) experience related to PRO completion before the visit; (3) experience related to the use of PRO results and the dashboard during the visit; (4) experience of any problems or disadvantages because of PRO; (5) comprehension, acceptability, and face validity of items and scoring algorithms; and (6) overall judgment, perceptions, attitudes, and recommendations regarding future use of the intervention.

All consultations and semistructured interviews were audio-recorded and transcribed verbatim. Four consultations were observed in person by a clinical diabetes psychologist for HCP supervision purposes to complement informant perspectives and assess any potential risks related to the way psychological issues are identified, addressed, followed up on, and reacted to.

Textbox 1. Evaluation questionnaires.

The Danish Patient-Reported Outcome (PRO) diabetes questionnaire (33-71 items)

- Completed by people with diabetes at home in advance of the visit. Evaluates general health and life situation, well-being, depression, symptom distress, annual check of feet and eyes, daily life with diabetes, worries about diabetes, confidence in managing diabetes, blood sugar regulation and hypoglycemia, medicine experience, access to health care professionals (HCPs), priority areas for self-management support, and preferred topics to focus on for the visit.

PRO diabetes questionnaire Evaluation Pilot Questionnaire for People With Diabetes (PRO-EVAL-P-1A; 4 items)

- Completed by people with diabetes immediately after the PRO questionnaire; evaluates perceptions of (1) relevance, (2) difficulty, (3) comprehension, (4) topic coverage/comprehensiveness, (5) acceptability, and (6) item-specific issues.

PRO Consultation Evaluation Pilot Questionnaire for People With Diabetes (PRO-CON-EVAL-P-1A; 14 items)

- Completed by people with diabetes immediately after the PRO questionnaire diabetes visit; evaluates perceived (1) support for autonomy and person-centered communication; (2) PROs impact on dialog, role, and care; (3) potential negative impacts; (4) face validity of scoring algorithms; (5) interest in continued use and advocacy; and (6) suggestions for improvement.

PRO Consultation Evaluation Pilot Questionnaire for HCP (PRO-CON-HCP-1A; 10 items)

- Completed by HCPs immediately after the PRO questionnaire diabetes visit; measures perceived (1) quality of visit; (2) PROs impact on dialog, roles, visit outcome; (3) challenges with the use of PRO; (4) satisfaction and interest in future use; and (5) clinical validity of items and algorithms.

Data Analysis

Quantitative data (blood tests, sociodemographic, PRO questionnaire, and Likert scale evaluation questionnaire) and qualitative data (consultation and interview transcripts, free-text evaluation responses, and notes from debriefing meetings) were

analyzed in SPSS Statistics 25.0 for Windows (IBM Cooperation) and NVIVO 12.0 (QSR International), respectively. Primary quantitative evaluation data were presented descriptively. The main qualitative analysis used a simple stepwise coding process adapted from thematic analysis and a phenomenological and combined inductive and deductive

approach [21-23]. The aim was to evaluate perceived feasibility, acceptability, benefits, and drawbacks among people with diabetes and HCPs while examining the extent to which PRO was perceived as having an impact on the participation of people with diabetes, quality of dialog, overall care visit, and follow-up. The steps used by the qualitative researcher included (1) all data being reviewed and iteratively coded using the research questions as topic structure; (2) all codes being analyzed for duplicity, relationships, and hierarchical structures; (3) key categories and themes being identified; (4) robustness of each category or theme being checked by reviewing each against all raw data; and (5) results being continuously checked and discussed with the multidisciplinary clinical team, multiple coders, and a panel of people with diabetes. The analysis included an exploratory element to assess whether benefits or disadvantages were highlighted, which did not fit within the predefined research questions. Consultation transcripts were analyzed using a predominantly semantic content analysis approach [24] to assess selected aspects of fidelity, feasibility, and acceptability regarding the use of PRO during the visit. A codebook was developed based on a review of all transcripts with 4 main coding categories: (1) HCP use of the PRO dashboard (sharing and use of open-ended prompts), (2) response of the person with diabetes when the HCP prompted a specific PRO result (validating its relevance), (3) follow-up

action in response to prompted PRO topics (eg, problem-solving dialog, referral, therapy change, education, and new appointment), and (4) PRO topic code (PRO item, scale, or construct). One researcher coded all transcripts to establish the codebook, and 3 students coded 3 interviews each to compare scoring and identify codebook ambiguities. Work meetings were conducted to refine the codebook until concordance in coding was established. Qualitative data were used for the explanatory analysis of the quantitative evaluation questionnaire data as a mixed-methods design.

Results

Overview

The characteristics of the 12 people with diabetes enrolled are shown in Table 1. A good variance was achieved in relation to gender, type of diabetes, age, duration of diabetes, therapy and treatment modality, complications, and comorbidity burden. The HCPs were 2 senior diabetes nurses and 2 senior diabetes physicians employed at the Ambulatory Diabetes Clinic at Aalborg University Hospital. All HCPs (3/4, 75% women and 1/4, 25% men) had >5 years of diabetes care experience, and all had had some previous involvement with the design of the PRO diabetes tool.

Table 1. Characteristics of study participants (N=12).

Characteristics	Values
Gender (female), n (%)	7 (58)
Type of diabetes, n (%)	
Type 1	8 (67)
Type 2	4 (33)
Age (years), median (range)	56.6 (24-79)
Duration (years), median (range)	19.5 (2-50)
Device used, n (%)	
Insulin pen	7 (58)
Insulin pump	2 (17)
GLP-1 ^a , n (%)	3 (17)
Tablet, n (%)	1 (8)
Number of comorbidities	
Mean	1.58
Median (range)	1 (0-8)
Number of complications	
Mean	1.3
Median (range)	0.5 (0-4)
Outcome variables, mean (SD; range)	
HbA _{1c} ^b (mmol/mol)	85.8 (20.7; 61-113)
Health (SF-1 ^c ; score range 1-5)	3.0 (0.7; 2-4)
Well-being (WHO-5 ^d ; score range 0-100)	60.4 (20.7; 20-96)
Diabetes-specific distress (PRO ^e Diabetes Questionnaire–Negative Impact of Diabetes Scale, three-item distress scale; score range 0-100)	41.0 (14.0; 8-58)
Number of PRO topics flagged for action (scored with a <i>red</i> flag)	3.8 (2.7; 0-10)
Number of PRO topics flagged for action (scored with a <i>yellow</i> flag)	4.8 (3.0; 1-11)
Total number of PRO topics flagged for action (yellow and red topics and additional topics selected by people with diabetes)	14.3 (6; 4-26)

^aGLP-1: glucagon-like peptide-1.

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

^cSF-1: global health item of the Short-Form Health Survey.

^dWHO-5: World Health Organization–Five Well-Being Index.

^ePRO: patient-reported outcome.

Quantitative Data Results

Evaluation of the PRO Diabetes Questionnaire by People With Diabetes

The results from the quantitative evaluation of the PRO questionnaire by people with diabetes are shown in [Table 2](#).

All the people with diabetes were positive about the relevance of the questions for their diabetes, except 1 person with diabetes who indicated a moderate negative appraisal. All people with diabetes found it easy to complete the entire questionnaire. Of the 12 participants, 6 (50%) indicated no items, and 6 (50%)

indicated *one or few items* of the PRO questionnaire were difficult to understand. The specific items were reviewed with people with diabetes in interviews as part of the formative questionnaire validation and design. Of the 12 participants, 10 (83%) indicated that no important topics were missing from the questionnaire, and 2 (17%) mentioned topics of *atypical diabetes type* and *more on how to access pump technology* as desired topics. Both interview and questionnaire data confirmed that people with diabetes felt the color-coded display of their PRO data, including specific cut-off thresholds in *DiaProfil*, provided a valid and helpful picture of their situation. Furthermore, HCPs provided separate confirmations of the validity of the scored PRO outputs for each person with diabetes after each visit.

Table 2. Quantitative evaluation of the patient-reported outcome diabetes questionnaire by people with diabetes (N=12)^a.

Response options	Value
How relevant did you find the questions to be for your diabetes care?	
Mean score (SD)	3.5 (0.7)
Negative appraisal^b, n (%)	
1	0 (0)
2	1 (8)
Positive appraisal^b, n (%)	
3	4 (33)
4	7 (67)
How difficult or easy was it for you to complete the questionnaire?	
Mean score (SD)	3.6 (0.7)
Negative appraisal^b, n (%)	
1	0 (0)
2	0 (0)
Positive appraisal^b, n (%)	
3	6 (50)
4	6 (50)
Were any items difficult to understand? n (%)	
No	9 (75)
Yes, one or a few	3 (25)
Yes	0 (0)
Did you miss any topics missing? n (%)	
No	10 (83)
Yes	2 (17)

^aShows the responses of people with diabetes to the PRO-EVAL-P-1A pilot evaluation questions immediately after completing the PRO questionnaire.

^bA score of 1 and 2 reflects a negative appraisal, and 3 and 4 represents a positive appraisal.

Questionnaire Evaluation of the Use of PRO During the Visit by People With Diabetes

The mean single-item scores by item for the primary evaluation questions are shown in [Table 3](#).

All 12 people with diabetes rated the person-centered autonomy-supportive communication style of the HCPs positively, with mean scores ranging between 4.5 and 4.8, with a score range of 1-5. People with diabetes felt that their HCPs were focused on their priorities, encouraged them to speak, and made them feel comfortable talking about their needs, and all felt they got the care and advice that they had hoped for. In order of decreasing positive rating, people with diabetes expressed high interest in continued use, that PRO should be

offered as standard care to all, that PRO helped focus on what was most important to them, that PRO helped focus the conversation on what mattered most to them, that they would like to use PRO in their future care, that they felt better prepared for the visit, and that the PRO dashboard provided a good picture of their current diabetes situation, needs, and priorities. There was only 1 person with diabetes who indicated a moderate degree of being uncomfortable or having a problem related to the use of PRO. During the interview, where answers were debriefed, she explained that she had had a bad day, was feeling very distressed because of diabetes, and had not felt the HCP understood her issues as they were raised. Despite this, she was very positive about the PRO tool and did not attribute the problem to the tool.

Table 3. Quantitative evaluation of the use of patient-reported outcome (PRO) during the visit by people with diabetes (N=12)^a.

Items	Single-item score ^b , mean (SD)
General dialog evaluation	
Did the HCP ^c make you feel comfortable talking about the topics that you needed to?	4.8 (0.6)
Did the HCP give you the treatment, advice, referral, or other assistance that you needed?	4.6 (0.5)
Did the HCP focus on what is most important to you?	4.5 (0.9)
Did the HCP encourage you to give input or ask questions?	4.5 (0.7)
Evaluation of influence of use of PRO	
Interested in using PRO in future care	4.9 (0.3)
PRO should be part of standard care	4.8 (0.4)
The <i>DiaProfil</i> dashboard helped me talk about the important things	4.7 (0.3)
HCP was more or less prepared because of use of PRO	4.4 (0.9)
I felt more or less prepared because of use of PRO	4.4 (0.9)
PRO dashboard gave a good picture of my situation, needs, and priorities related to my diabetes	4.3 (1.0)
Any problems or uncomfortable experiences related to the use of <i>DiaProfil</i> in the dialog? If so, which?	1.3 (0.6)

^aMean single-item scores of evaluation by people with diabetes of use of PRO after the visit using the PRO-CON-EVAL-P-1A-pilot questionnaire.

^bScore range is 1-5, except for the last item, which is 1-3.

^cHCP: health care professional.

Results of the Qualitative Analysis of Interviews With People With Diabetes

The main themes identified from the qualitative analysis of the interviews are shown in [Table 4](#), with illustrative quotes.

People with diabetes expressed that they felt the PRO questionnaire covered all relevant issues, was straightforward to fill out, and facilitated positive self-reflection. In line with the questionnaire evaluations, approximately half of the people with diabetes reported minor issues with understanding one or a few items.

Most people with diabetes expressed that they found it positive to complete the questionnaire at home in advance, as it helped them know what the conversation would be about at the visit. Filling out the questionnaire made the people with diabetes feel reassured that they would remember and get to talk about their priority issues with their HCP. This was important as several people with diabetes expressed frustration that they would often forget to talk about the issues that mattered to them during visits. A family member who participated with a study participant explained the following:

It is a nice thing to fill it out at home, and the HCP is also better prepared for what you want to ask about. I would like to have this every time.

Several people with diabetes had a positive personal experience filling out the questionnaire. One person with diabetes said the following:

When I sat down with the questionnaire, I had some time for it, and I felt really positive and surprised. Because it went straight in and touched on some issues where I had to feel and actually remove the shutters and relate to it. "How am I doing? If I am totally honest, how is it going with this?" "It is easy to say, "well it's going to be fine." "Let's just keep going as usual. So, in this way it was really an eye-opener for me.

Only one person with diabetes expressed an uncomfortable situation related to the use of PRO during the visit, as noted earlier, and it was clarified in the interview that the person with diabetes did not attribute the issue to the PRO tool but attributed it to not feeling fully understood by the HCP.

Table 4. Analysis of semistructured interviews with people with diabetes.

Theme and subcategories	Quotes for illustration
Perception of the questionnaire	
Easy and manageable to fill out	
<ul style="list-style-type: none"> • Easy to do • A positive experience • No or few items difficult to understand 	“I felt it was easy and I think it was well laid out with 5 response options every time—and I just felt it was simple”
Relevant comprehensive questionnaire	
<ul style="list-style-type: none"> • All questions relevant • No key topics missing • 360° coverage is good 	“It gets into all the issues, which I think is good because it makes you think about stuff you might not otherwise have”
Benefits related to filling out the questionnaire	
Feel better prepared	
<ul style="list-style-type: none"> • Self-reflection and self-insight 	“It motivated me hugely to get some things on the table that I have been needing to talk about but closed my eyes to because I tend to just say things are 'kind of fine'?”
<ul style="list-style-type: none"> • Comfort knowing own priorities will be addressed in the visit 	“Many times I feel I forget when I get to the visit, shoot, this or that I forgot to ask about when you sit there—and when you leave you haven't asked about what you needed. So, this is really super this tool”
Experience of the process of use of PRO^a in-visit	
HCP^b uses PRO in pleasant or good way	
<ul style="list-style-type: none"> • PRO screen was easy and intuitive in visit 	“I think it was great to see [DiaProfil screen]—it made sense; red is bad, yellow is less bad and green that is perfect-ish, right?—I think it gave a good picture [of my situation] and it was easy for me to grasp it”
<ul style="list-style-type: none"> • HCP talked about PRO results in a pleasant way 	“I liked the way that she [HCP] went at it right away. She made me actually feel reassured by showing me the screen and just mentioning e.g. there was this box with a mental issue”
Benefits of PRO during the visit	
Better, more meaningful care visit	
<ul style="list-style-type: none"> • HCP is better prepared 	“It felt very nice [that the nurse had seen my PRO answers], because I am thinking if she has read it through she might see some connections between my issues – makes sense.”
<ul style="list-style-type: none"> • Cover new important topics that matter to me 	“We covered some other topics than normal, because I usually just get the numbers [blood test results] and sometimes gets measured and weighted and then go home again. It got more personal. And I think that was awesome.”
<ul style="list-style-type: none"> • Easier to talk about difficult or sensitive issues 	“It was a bit easier to sit at home and write that I actually would like to get some help to stop smoking than to sit down here—now it's out in the open without having to say it face to face. It also makes it easier for the HCP I think—you can maybe get to talk about some of the difficult topics that you wouldn't just sit there and say.”
<ul style="list-style-type: none"> • Getting better care 	“I get better help by answering these questions”

^aPRO: patient-reported outcome.

^bHCP: health care professional.

HCP's Evaluation of the Use of PRO During the Visit

The main results from the HCP questionnaire evaluation of the visits are shown in [Table 5](#). HCPs rated the general quality of the dialog positively, felt they were able to cover all the issues that were important to them, and were satisfied with the clinical outcome. In one case, the HCPs noted general dialog difficulties; however, this was reported by the HCP to be unrelated to PRO and caused by the fact that the person with diabetes had mild cognitive impairment and language difficulties. In order of

decreasing positive ratings, HCPs were satisfied with the use of PRO, highly interested in continued use of the tool, and felt that all relevant topics were covered by the PRO questionnaire. HCPs reported moderate improvement in preparedness and active engagement among people with diabetes during the visit. HCPs were neutral or marginally positive about the ability of PRO to uncover previously unknown clinical challenges of people with diabetes, and in 50% (6/12) of visits, HCPs highlighted some challenges related to the use of PRO during the visit.

Table 5. Results of questionnaire evaluation of the use of the patient-reported outcome (PRO) diabetes tool by health care professionals (N=12)^a.

Item content ^b	Mean single-item score (SD)
General dialog evaluation	
How would you rate the overall quality of the dialogue with this patient?	4.25 (1.1)
Did you cover all the topics in the visit that were important to you from a clinical perspective?	4.4 (0.7)
Evaluation of influence of use of PRO	
How interested are you in using <i>DiaProfil</i> in its present form in your future diabetes consultations?	9.25 (1.0)
How satisfied or dissatisfied were you overall with use of <i>DiaProfil</i> in the consultation?	4.6 (0.7)
Was the person with diabetes was more or less prepared for your dialog due to answering the PRO questionnaire?	3.75 (0.75)
Do you feel that your use of <i>DiaProfil</i> changed your role in the conversation?	3.5 (1.0)
Did you experience that the person with diabetes got to speak more or less during this visit due to use of <i>DiaProfil</i> ?	3.6 (0.5)
Did <i>DiaProfil</i> make you aware of clinically relevant issues for this patient you were not aware of before?	2.67 (1.3)
Did you experience challenges due to the use of <i>DiaProfil</i> during the visit?	2.33 (1.4)

^aMean single-item scores of health care professional evaluations of use of patient-reported outcome in the visit (PRO-CON-EVAL-P-1A-pilot).

^bScore range is 1-5. Score of 1-2: negative; 3: neutral; and 4-5: positive (except for the question on interest in continued use, which has a score range of 1-10).

Results of the Analysis of HCPs' Free-Text Evaluations of Use of PRO in the Visit

Themes and illustrative quotes from the HCP's evaluation of the visits are shown in Table 6. HCPs felt that PRO helped focus the dialog during most visits, and it was mentioned that the fact that people with diabetes had reflected on issues in advance contributed to this. Generally, HCPs did not indicate that PRO revealed new clinical treatment insights; however, in half of the visits, PRO placed attention on the fact that clinically relevant issues were of specific importance to their patients, with sleeping difficulty, erectile dysfunction, cardiovascular symptoms, worry about late-stage complications, foot problems, and psychological issues being problems specifically mentioned. The PRO tool was noted to provide structured insight, especially into the

psychological factors at play. HCPs did not report any technical or functional issues related to using the digital tool but reported a number of challenges in half of the visits, reflecting in large part the fact that the HCPs were new to using the tool in a routine visit and there was no detailed manual for combining PRO and clinical data during the visit. Key issues raised by HCPs were how to handle the dialog if the person with diabetes had a lot of PRO topics requiring attention and that it could be difficult to find the right balance between traditional clinical tasks and the new PRO topics. Another issue that was raised was that some symptoms flagged in the PRO dashboard were not always related to diabetes or had already been acted upon by another HCP, so the value of raising these issues appeared unclear to the HCP.

Table 6. Analysis of open-ended text evaluations by the health care professional after each visit.

Main themes	Quote or case example
Benefits of use of PRO^a in the visit	
The person with diabetes was better prepared due to self-reflection in advance	“I tend to ask a lot of questions if the patient does not say so much. In this case the patient had already reflected and prioritized which allowed us to focus on this instead of ‘shooting in the blind.’”
Better able to set agenda in line with patient priorities	“I distributed the available time better, focused on the problems of the patient, listened more”
New insights about which topics are important to the person with diabetes.	New topics identified as important to the person with diabetes included worry about complications, foot wound, erectile dysfunction, sleep, pain, and barriers in life situation to managing diabetes.
Challenges regarding use of PRO in the visit	
Managing the conversation when there are many flagged PRO topics	“Dialogue would have been better if we had had more time to wrap up the various problem areas identified” “There were so many answers to consider that it extended the visit.”
Not all PRO outputs were relevant to act on	
Red score on pain	A person with diabetes scored red on pain, but it was because of arthritis pain that was already treated and addressed in other care setting.
Red score on low well-being	A person with diabetes scored red on low well-being, but it was because of life issues unrelated to diabetes.
Uncertainties regarding use	Unsure how to handle a discrepancy between a PRO score and what the person with diabetes says in the visit.

^aPRO: patient-reported outcome.

Content Analysis of Transcripts of Diabetes Visits

The results of the content analysis of the audio recordings of the visits are summarized in [Table 7](#). HCPs generally used the PRO dashboard as intended for dialog support (albeit in different ways and to different degrees). As there was no video recording of the visits, we could not make precise assessments of when and how the PRO dashboard was used to prompt each individual topic. This was especially difficult as both the person with diabetes and HCP would view the screen while talking. Overall, with few exceptions yellow or red colored topics and topics requested by the person with diabetes on the PRO dashboard questionnaire were prompted or acted upon during the visit by the HCP. Several topics were only briefly mentioned; however, they were reviewed and considered for relevance. The key topics that were shown to be both flagged as a potential problem area on the PRO dashboard, prompted in the visit by the HCP, confirmed as relevant by the person with diabetes and followed-up on by the HCP are depicted in [Textbox 2](#).

Many PRO topics prompted by the HCPs were validated as relevant by the person with diabetes and, for the most part, resulted in relevant follow-up dialog and, in some cases, action. However, in line with HCP evaluations, there were several instances where topics raised pertaining to, for example, symptoms unrelated to diabetes were found not to lead to concrete action or follow-up plans. By cross-matching the topics discussed during the visit with the PRO data, we identified some individual errors because of mistakes by people with diabetes during questionnaire completion; however, we could exclude any structural PRO assessment problems and could confirm the clinical validity and utility of the PRO outputs. Observations of a subset of consultations by a clinical diabetes psychologist provided additional reassurance that it was possible for HCPs to incorporate PRO data in a person-centered manner into the dialog. Observations provided input to our future person-centered training for the use of PRO, especially regarding identifying and clarifying previously undetected psychosocial problems.

Table 7. Results of analysis of audio-recording transcripts of patient-reported outcome (PRO) diabetes consultations (N=12).

Category	Case	Result, n (%)
Fidelity of in-visit use of PRO		
HCP ^a used open-ended questions to prompt people with diabetes about a flagged PRO topic or result	“Is this something that you recognize?” [HCP] ”Here it says something about pain in your feet?” [HCP]	12 (100)
HCP showed and explained the PRO dashboard	“What you answered in the questionnaire is shown on this screen.” [HCP] ”If it is green it is all good, like a traffic signal, right?” [HCP] ”It is like a traffic light. Yellow is something to be attentive to if there might be something to talk about.” [HCP]	10 (83)
HCP explicitly used a PRO to set visit agenda	“And then there is red which we definitely should talk about...” [HCP] Case example: “Is it things that bother you?” [HCP]; “Yes” [person with diabetes]; “Ok then it makes sense we start out with this” [HCP]	8 (67)
All flagged PRO topics were mentioned during the visit.	“You have indicated you feel that diabetes takes up too much of your daily life?” [HCP]	12 (100)
Use of PRO in the visit		
At least one PRO topic was prompted by the HCP, validated as relevant by the person with diabetes, and followed up with action ^b	Case example: PRO topic flagged: “Not confident in ability to get contact with HCP if needed”. HCP is prompted with open question. Person with diabetes validates result. HCP is informed about contact options and handed out a diabetes telephone hotline leaflet.	12 (100)

^aHCP: health care professional.

^bActions may include follow-up dialog, advise, treatment, education, self-help resources, care planning, and referrals.

Textbox 2. Topics where cases of intended clinical use of the patient-reported outcome tool were identified in this study.

Topics
<ul style="list-style-type: none"> • Life situation impacting diabetes • Psychological well-being • Daily life with diabetes • Worry about complications • Blood sugar regulation and hypo/hyperglycemia worry • Diet, carbohydrate counting, and smoking • Medicine experience • Symptom distress: Sleep, sexual function, foot, neuropathic pain, gastrointestinal, and cardiovascular • Confidence in access to health care professionals for diabetes

Mixed-Methods Analysis

Overall Results

Qualitative and PRO data were compared with questionnaire evaluation data to verify the findings and establish robust overall findings regarding feasibility, acceptability, and perceived benefits. Both questionnaire and qualitative data confirmed that the PRO questionnaire and the *DiaProfil* tool were acceptable, valid, appropriate, and effective in improving people with diabetes' preparation for the visit and the quality of the dialog.

Both questionnaire data and interviews for people with diabetes and visit transcripts supported that HCPs used the PRO dashboard in the intended person-centered way and achieved a more focused and relevant dialog. The HCP questionnaire and open-ended text confirmed general satisfaction with the use of PRO tools.

Some people with diabetes had a high number of flagged PRO topics in this version of the tool, and some PRO-flagged topics were not relevant for the HCP to act on in the specific visit; however, people with diabetes were satisfied with the use and did not report concerns or negative experiences related to this. We used a multi-informant mixed-methods approach to analyze negative outliers in the questionnaire data. As an example, we looked at the only person with diabetes who had rated the relevance of the PRO diabetes questionnaire as somewhat not relevant, whereas the 11 others had rated it relevant. The person with diabetes had had diabetes for >13 years, was resourceful and empowered in relation to diabetes, and had seen the same physician for several years. The person with diabetes's *DiaProfil* dashboard was largely *green* but with flagging of pain, sleep, and blood sugar measurement. She indicated that as she was doing fine and knew what she wanted to talk about; she did not find the questions so relevant when filling out. The person with

diabetes expressed afterward that “we covered other topics than usual which made it more personal—I think it was great” and was keen to keep using it. Another point the person with diabetes raised was that a question about medicine required for branch logic was irrelevant as the hospital should already have the information, which was true, and we began to prepopulate that item subsequently. This highlights the significance of iteratively testing every single item with users. The HCP evaluation confirmed that the person with diabetes was resourceful and knew what they wanted from the visit; the HCP also answered neutral (3) on whether the person with diabetes was better or worse prepared because of PRO.

Evaluation of the Acceptability of the PRO Evaluation Questionnaires

People with diabetes were able to fill out the evaluation questionnaires digitally without the need for support after each visit and expressed, in interviews, that the questions were relevant and easy to understand and fill out. We obtained complete evaluation data from both people with diabetes and HCP from all visits. This corroborated our initial finding from user-testing that they were suitable for use in routine care situations. As part of the formative design, the pilot evaluation questionnaires were revised based on our mixed-methods analysis and feedback from people with diabetes for use in a real-world study of how people with diabetes experience use of PRO [25]. Changes included alignment of items for people with diabetes and HCPs to evaluate interrater reliability, additional questions relating to perceived impacts of PRO on self-management versus HCP decision-making, role of the use of screens during the visit, emotional impact of PRO, and impacts of PRO on HCP’s work satisfaction and work stress. The final questionnaires were incorporated into a large-scale real-world study of the PRO diabetes tool in different health care settings [25].

Discussion

Principal Findings

The PRO tool helped improve the quality of the dialog and visit by facilitating the identification and prioritization of topics according to the needs of people with diabetes. In line with the extant PRO literature, the use of PRO helped introduce new talk topics, supported dialog about difficult psychosocial issues previously left unattended, and facilitated a more active role for people with diabetes [3,26,27]. People with diabetes felt better prepared for the visit, and in several cases, HCPs reported that people with diabetes appeared more ready to talk about difficult but clinically relevant issues, which has also been reported by other studies. The use of the PRO tool was feasible, acceptable, and facilitated a person-centered dialog in line with findings from other studies regarding the use of PRO to facilitate aspects of person-centered care [2,3,26-30].

Meeting the Expectations of the People With Diabetes and Family Members

A common concern expressed by HCPs is that PRO questionnaires may raise unrealistic expectations for people with diabetes and that resources are not available to address

topics raised by PRO [26,31,32], which could result in disappointment during the visit. All people with diabetes expressed that they received the care or support they needed from the visit; however, 1 person with diabetes expressed that the HCP did not sufficiently understand the person’s issues related to diabetes distress. All 12 people with diabetes in our study were positive about the PRO tool and the way the HCPs used it and did not report negative effects that they felt should prevent it from being a part of standard care. Although our combined analysis identified some issues identified by PRO that were not found relevant to act on, people with diabetes did not report any concerns related to lack of follow-up on PRO results or otherwise unfulfilled expectations.

Our finding that none of the people with diabetes expressed disappointment or unfulfilled expectations may be partly because the questionnaire had been specifically designed to only include items that were perceived as directly relevant for routine care by both people with diabetes and HCPs. Given the importance of expectation setting in the clinical use of PRO, we attempted to give instructions to people with diabetes about PRO carefully to avoid inadvertently raising unrealistic expectations.

Fidelity and Use of Person-Centered Strategies

Our combined analysis indicated good fidelity related to the HCP’s use of the basic recommendations for person-centered use pertaining to the use of open questions to verify and clarify priorities of people with diabetes and ensure coverage of all key topics highlighted by them. The use of open questions and active listening was applied in all visits, and we found it to be an essential component that may be particularly important in the care of people with cognitive or language difficulties who may have difficulties completing the PRO questionnaire as intended. The PRO intervention did not include requirements for HCPs to use specific methods or tools for agenda-setting, structuring the visit, collaborative goal-setting, action-planning, and shared decision-making beyond the general recommendations for person-centered use of the tool. Therefore, this study provided a first opportunity for HCPs to try out the PRO tool in routine visits, identify key communication challenges, and begin to develop individual strategies on how to use the tool in an optimal way. The main challenges reported by the HCPs relate to how to structure the time and balance focus on clinical and PRO issues. An HCP expressed that it could be difficult to juggle tasks during the visit when people with diabetes had a high number of PRO topics, and an HCP found it difficult to review the many topics on the screen while maintaining a natural conversation. In contrast, and interestingly, the people with diabetes did not express concerns or problems related to the number of topics being identified and were overwhelmingly positive about the way the HCPs used the tool. We believe this may be partly because of the HCPs putting in an extra effort to ensure that people with diabetes had a good dialog experience despite the high number of flagged issues.

In a few cases, it was noted by HCPs that some flagged PRO topics, especially those related to symptoms, were not related to diabetes or were not relevant to act on in this visit; however, it was still necessary to review the topics as they were flagged on the PRO dashboard. The issue of not being able to act on

especially generic PRO issues has also been reported by others [33]. Importantly, we found that people with diabetes appreciated the dialog on these topics, such as pain, sexual dysfunction, and sleep issues, as these issues were very distressing to them. Since the study, the PRO tool was further adjusted to include comorbidity information to facilitate the identification of symptoms unrelated to diabetes. Our results highlighted that the use of the PRO tool imposes significant demands on HCPs in relation to person-centered communication skills, self-efficacy for the use of PRO, and knowledge, as also highlighted by others [15,34]. Although the PRO tool was designed to be usable by diabetes HCPs with only minimal PRO-specific training, we believe that sparring with senior colleagues in developing individual person-centered strategies for the use of the PRO tool that suits each HCP's style and clinical responsibilities will be important for optimal implementation and effectiveness. Clarifying each HCP team member's role and responsibility in relation to the key PRO topics using frameworks such as 5A [35] may be important to support HCP adoption and self-efficacy for PRO use. PRO tools differ greatly in terms of how prescriptive the guidance is on how to apply PRO data for goal-setting and action-planning [30,36]. Further research is needed and underway to examine how different person-centered styles and communication strategies impact the use and impact of PRO [3,4,37,38].

Benefits Related to Completion of the Questionnaire

People with diabetes expressed positive experiences related to the completion of the questionnaire. We found several factors that could potentially explain this. The completion of the questionnaire led to a reflective process that facilitated self- and disease-insight related to diabetes. During the design of the PRO tool, people with type 2 diabetes highlighted the potential for this questionnaire to help people with diabetes understand their situation and options for acting for their own health. On the basis of our results, this effect is likely important for a subgroup and dependent on diabetes duration, empowerment, and whether the person with diabetes is completing the PRO questionnaire for the first time. Filling out the questionnaire at home made people with diabetes feel confident that they would get to talk about their priority issues during the visit. This was very important as people with diabetes reported that they do not normally prepare for visits and often forget to ask about issues that are important to them. Furthermore, this appeared to have a potential impact on expectations and motivation related to participating in the upcoming visit.

Benefits of a Broad Topic Coverage of the PRO Diabetes Questionnaire

The broad and balanced coverage of topics in the PRO diabetes questionnaire represents the result of an extensive iterative national multistakeholder participatory design process to achieve a core questionnaire acceptable and useful for both people with diabetes and HCPs in different care settings [3]. Our combined analysis found that a broad range of topics was relevant, usable, and perceived as beneficial either by the person with diabetes or the HCP. In addition to topics related to general life issues affecting diabetes, daily life with diabetes, self-management, and blood sugar regulation issues, people with diabetes

expressed appreciation for the opportunity to speak about symptom distress related to pain and to address distress because of sexual dysfunction, which are topics not included in, for example, depression and diabetes stress-specific PRO tools [39,40]. Different PRO constructs have been shown to be suitable for different clinical purposes, such as dialog support, treatment decision support, screening, symptom monitoring, outcome evaluation, and collaborative care planning [12,27,41]. Some PRO constructs are mainly used for screening, such as hypoglycemia unawareness [42,43], cardiovascular symptoms [44], and depression [45], whereas items about, for example, medicine experience may be particularly useful for treatment monitoring and support [46,47]. Items about individual priorities and confidence in self-management can facilitate behavior change, collaborative care, and individualized self-management support [34,48,49].

Our study supports that it is feasible to use a PRO diabetes questionnaire with a very broad range of topics and that the questionnaire may generate both specific benefits related to each PRO construct as well as related to its use as an overview to facilitate a person-centered dialog [14]. We believe that the high acceptability and satisfaction with the tool by people with diabetes across diabetes type, age ranges, and different levels of disease progression may be partly because of the broad topic coverage, broad response options, and the use of branching logic. Although it was a focus area for the project if the questionnaire was burdensome, most people with diabetes appreciated the comprehensiveness of the questionnaire and expressed satisfaction with the experience of completing it as it *touches all the bases*. People with diabetes and HCPs did not report that any major topics were missing or that there were topics that were irrelevant. People with diabetes and HCPs confirmed that the scoring algorithms applied for the color-coding responses were suitable and fit for the purpose.

Further research is warranted to investigate the importance of a broad topic coverage to facilitate the potential therapeutic and empowering effect of self-completing the PRO. It is possible that PRO instruments that focus only on one or a few topics may not provide the same support for disease insight, care navigation, and active participation as instruments which provide a comprehensive 360°-review of diabetes issues

Another possible benefit related to the broad topic coverage, when compared with, for example, depression-focused diabetes screeners, is that people with diabetes who may be doing well emotionally are still given an opportunity to express other priorities and issues that affect them in relation to diabetes.

HCPs noted that an important benefit of the PRO tool was related to obtaining a structured overview of each person with diabetes' overall situation, which was made possible by the comprehensive topic coverage. In a previous study, we found that there is a need to use a broad set of PRO outcome constructs to evaluate outcomes in diabetes as the needs of people with diabetes vary individually and change over time [16]. Recent standards for person-centered diabetes care also emphasize the significance of a comprehensive whole-person orientated evaluation of individual needs, preferences, and priorities [50].

Our PRO tool includes 5 items specifically related to depression and diabetes-related distress as it is important to legitimize and prompt dialog on these often insufficiently addressed issues [51,52]. Balancing items regarding these topics with items that reference daily life, self-management, symptoms, blood sugar regulation, medicine, and access to care may be important to set balanced expectations for visits. Although only an important minority of people with diabetes have mental health or emotional problems [53,54], it may be advantageous to invite people with diabetes to share their perspectives on all key aspects of care rather than only to express emotional symptoms.

Initial Insights Regarding Evaluating the Public Health Impact Potential of the PRO Tool

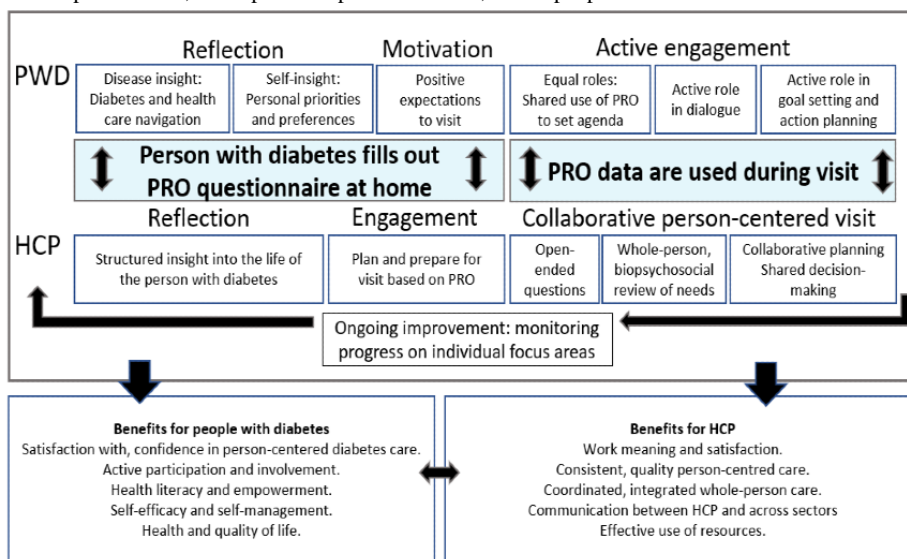
This study provides initial insights on how to evaluate the potential public health impact of the PRO tool using the dimensions of RE-AIM (reach, efficacy, adoption, implementation, and maintenance) [9,55-57]. In terms of reach, all people with diabetes found the PRO tool usable and useful, which may reflect the fact that a large heterogeneous group of people with diabetes was involved in all stages of its development. In terms of efficacy, our study suggests that the tool has the potential to improve the experience of care and

quality of diabetes visits in relation to person-centered principles. Future research is underway to examine the implications of the use of the tool for care quality, health, and cost outcomes. Regarding HCP adoption of the tool, HCPs were satisfied and interested in continuing the use of *DiaProfil*, and key barriers were found that were addressed by tool design and facilitation support. Further research is required to identify and address perceived barriers and facilitators to HCP adoption in a larger group of HCPs reflecting different disciplines and health care settings [14,25,58,59]. Implementation with adequate fidelity was established as possible by both physicians and nurses in our study. In terms of the PRO tool’s potential to become a permanent part of routine practice, that is, maintenance, we were encouraged by all people with diabetes expressing they felt that the PRO tool should be a part of standard of care without the need for further development.

Overall Implications and a Conceptual Model

Figure 3 shows an initial conceptual model to illustrate the hypothesized facilitators, barriers, and impacts related to the use of the PRO tool based on our combined data. This model will be expanded and improved as further and broader clinical and exploratory research is undertaken to incorporate a broader range of experiences, facilitators, barriers, and impacts.

Figure 3. Conceptual working model for hypothesized processes and impacts for the use of the patient-reported outcome diabetes tool in a routine diabetes visit. HCP: health care professional; PRO: patient-reported outcome; PWD: people with diabetes.



The first stage of reflection and improved disease- and self-insight and motivation facilitated by the questionnaire completion by people with diabetes were found to be important to the study participants and suggest the potential for the intervention to facilitate diabetes-related empowerment [4,5,20].

Both people with diabetes and HCPs reported that people with diabetes were more actively engaged as a result of PRO. In this pilot study, HCPs used the PRO tool without detailed protocolized steps for agenda-setting, shared decision-making [60], goal-setting, or action-planning [61], so the visits involved a learning experience where they each identified relevant strategies to use.

The anticipated beneficial impact of the PRO tool for people with diabetes relates to confidence in diabetes management

[62], self-efficacy [63] and empowerment [4,5,20,64], self-care behaviors [65], improvements to health and diabetes-related outcomes [16], and broader humanistic and societal outcomes [66] resulting from derived impacts.

HCPs appreciated gaining a structured insight into the lived experience of diabetes among people with diabetes, which facilitated reflection and allowed them to provide more personalized care. The PRO tool facilitated the introduction of daily life and psychosocial issues to achieve a comprehensive biopsychosocial review of the person with diabetes’ situation. As HCPs achieve self-efficacy for the use of PRO with their patients, there is a potential to experience improved work satisfaction and fulfillment. Given the utility of a wide range of topics, we believe the PRO tool has a particularly high

potential for improving the individualization of treatments and use of a wider set of support resources in accordance with what can benefit individuals with diabetes the most, thereby potentially promoting better self-management, outcomes, and use of resources.

The *DiaProfil* tool was specifically designed to facilitate person-centered chronic illness care [67,68] and enable coordinated self-management and psychosocial support for people with diabetes across health settings. As part of the formative evaluation, we found that the PRO tool is helpful in raising awareness and use of a broader range of services such as diabetes education, social services, and other specialist care. Although the HCPs in this study used open-ended questions when using the PRO tool, this may not always be the case. Additional research is needed to evaluate how variations in the use of person-centered communication by HCP may impact the effectiveness of PRO and whether there may be differential impacts on vulnerable patient groups affected by factors such as low health literacy, low socioeconomic status, and social and psychological challenges [15]. Further research is specifically needed to examine the mechanisms for and potential benefits of using the PRO tool to strengthen navigation support for people with diabetes and coordination of care across sectors [69].

An important additional aim of the PRO diabetes tool is to improve care by monitoring outcomes that matter to people with diabetes for value-based care [16,25,70]. During the formative evaluation of the PRO tool, people with diabetes have been enthusiastic about the prospect of using the PRO tool evaluation to monitor progress and changes over time, and the functionality is built-in. Follow-up studies are underway to examine how and to what extent the use of the PRO data as outcomes can facilitate improvements in care.

Limitations and Strengths

Our results should be examined while considering that this was a formative pilot study of the first viable version of the digital PRO tool, *DiaProfil*, before it was fully finalized for clinical use. This pilot study has important limitations: the study was not designed to provide empirical evidence for clinical effectiveness, and there was no attention control group, so it is not possible to rule out bias because of social desirability or effects related to study participation and added attention [71].

The small number of people with diabetes and HCPs limits the basis for generalization; even so, purposive sampling provided a good basis for examining experiences from a group of people with diabetes, which was diverse in terms of age, gender, type of diabetes, duration, treatment, and disease progression.

The 4 HCPs were previously involved in the design of the PRO tool, which limits the generalizability of their experiences.

Further research is required to evaluate the adoption and implementation by a diverse group of PRO-naïve HCPs in different health care settings.

Despite these limitations, our use of purposive sampling, multi-informant, and mixed-methods data analysis provided an opportunity to show the robustness of our core findings. Further research is required and planned based on the detailed findings in the study to examine the impacts of, facilitators of, and barriers to effective and integrated standard use of our PRO diabetes tool on a larger scale in different health care settings [25]. Additional questions incorporated in our future study include questions relating to the perceived benefits of PRO among people with diabetes for self-management, care quality, and treatment outcomes, and HCPs' confidence and skills in use of PRO in a person-centered manner and perceptions of impact of PRO on stress and work satisfaction.

Conclusions

This is the first study to show the feasibility, acceptability, and perceived benefits of using the Danish PRO diabetes questionnaire and the *DiaProfil* tool in a routine diabetes care setting. We found that people with type 1 and type 2 diabetes in different age groups found the digital PRO diabetes questionnaire relevant, acceptable, easy to use, and as having good coverage of relevant topics. Both people with diabetes and HCPs found the digital PRO tool, *DiaProfil*, feasible, appropriate, and value-adding for use in routine diabetes care. The PRO tool helped improve the preparation and active engagement of people with diabetes and improved the quality of the dialog in line with the objectives of the PRO tool. Filling out the PRO questionnaire had a positive effect on a subgroup of people with diabetes by facilitating self-reflection and better visit preparation, which also contributed to readiness for talking about needed but difficult issues with the HCPs. HCPs were able to use the PRO tool in a person-centered manner and were satisfied with its functionality. Minor issues and challenges were identified, which were addressed as part of the participatory development process for the PRO tool related to some people with diabetes having a high number of PRO topics. Adjustments were made to the PRO tool to address these issues.

In conclusion, we found that our newly developed PRO diabetes tool, consisting of the national PRO diabetes questionnaire and the digital PRO tool *DiaProfil*, was acceptable to people with diabetes and HCPs, provided clinically useful dialog and decision support to HCPs, facilitated active participation of people with diabetes, and overall improved the perceived quality of diabetes care visits. Further large-scale, comparative, and controlled research is now warranted to examine its potential for large-scale implementation and positive public health impact.

Acknowledgments

Members of the authors' user panel consisting of people with type 1 and type 2 diabetes provided invaluable input and codeveloped study materials along the way: Dorthe Hinzman, David Rasmussen, Henning Nielsen, Anni Fynbo, and Tove Brix.

The authors would like to thank the Department of Endocrinology and Steno Diabetes Center North Denmark at Aalborg University Hospital for invaluable support for participatory design, user-testing, study recruitment, and data collection; Karleen Parquette for observing visits and providing feedback regarding the health care professional's handling of psychosocial issues identified by patient-reported outcome; and Lise Boel and medical students Mikkel Wandahl Mathiesen, Sebastian Lawrence, and Frederik Østergaard Klit for providing support for the transcription of interviews, consultations, and coding. The region of North Denmark provided financial support to make this research possible. The authors would also like to thank Zitelab as their partner in the participatory design of the digital *DiaProfil* tool and data collection.

The pilot evaluation questionnaires and study materials were designed by the lead author (SES) using qualitative research and iterative development with input from people with diabetes and health care professionals. They were used to develop the final PRO-EVAL and PRO-CON-EVAL questionnaires which are now used in real-world studies of patient-reported outcome also in other disease areas. The final evaluation questionnaires can be obtained for research use without license costs by contacting the author.

Conflicts of Interest

None declared.

References

1. Bootle S, Skovlund SE, Editorial Group of the 5th DAWN International Summit. Proceedings of the 5th International Dawn Summit 2014: acting together to make person-centred diabetes care a reality. *Diabetes Res Clin Pract* 2015 Jul;109(1):6-18 [FREE Full text] [doi: [10.1016/j.diabres.2015.04.004](https://doi.org/10.1016/j.diabres.2015.04.004)] [Medline: [25979275](https://pubmed.ncbi.nlm.nih.gov/25979275/)]
2. Dickinson WP, Dickinson LM, Jortberg BT, Hessler DM, Fernald DH, Cuffney M, et al. A cluster randomized trial comparing strategies for translating self-management support into primary care practices. *J Am Board Fam Med* 2019 May 08;32(3):341-352 [FREE Full text] [doi: [10.3122/jabfm.2019.03.180254](https://doi.org/10.3122/jabfm.2019.03.180254)] [Medline: [31068398](https://pubmed.ncbi.nlm.nih.gov/31068398/)]
3. Skovlund SE, Lichtenberg T, Hessler D, Ejlskjær N. Can the routine use of patient-reported outcome measures improve the delivery of person-centered diabetes care? A review of recent developments and a case study. *Curr Diab Rep* 2019 Aug 16;19(9):84. [doi: [10.1007/s11892-019-1190-x](https://doi.org/10.1007/s11892-019-1190-x)] [Medline: [31420754](https://pubmed.ncbi.nlm.nih.gov/31420754/)]
4. Funnell M, Weiss M. Empowering patients with diabetes. *Nursing* 2009 Mar;39(3):7. [doi: [10.1097/01.nurse.0000347073.76891.2a](https://doi.org/10.1097/01.nurse.0000347073.76891.2a)]
5. Gómez-Velasco DV, Almeda-Valdes P, Martagón AJ, Galán-Ramírez GA, Aguilar Salinas CA. Empowerment of patients with type 2 diabetes: current perspectives. *Diabetes Metab Syndr Obes Targets Ther* 2019 Aug;Volume 12:1311-1321. [doi: [10.2147/dmso.s174910](https://doi.org/10.2147/dmso.s174910)]
6. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012 Oct 3;21(8):1305-1314. [doi: [10.1007/s11136-011-0054-x](https://doi.org/10.1007/s11136-011-0054-x)] [Medline: [22048932](https://pubmed.ncbi.nlm.nih.gov/22048932/)]
7. Adu MD, Malabu UH, Callander EJ, Malau-Aduli AE, Malau-Aduli BS. Considerations for the development of mobile phone apps to support diabetes self-management: systematic review. *JMIR Mhealth Uhealth* 2018 Jun 21;6(6):e10115 [FREE Full text] [doi: [10.2196/10115](https://doi.org/10.2196/10115)] [Medline: [29929949](https://pubmed.ncbi.nlm.nih.gov/29929949/)]
8. Eikelenboom N, van Lieshout J, Jacobs A, Verhulst F, Lacroix J, van Halteren A, et al. Effectiveness of personalised support for self-management in primary care: a cluster randomised controlled trial. *Br J Gen Pract* 2016 Apr 14;66(646):354-361. [doi: [10.3399/bjgp16x684985](https://doi.org/10.3399/bjgp16x684985)]
9. Glasgow RE, Dickinson P, Fisher L, Christiansen S, Toobert DJ, Bender BG, et al. Use of re-aim to develop a multi-media facilitation tool for the patient-centered medical home. *Implement Sci* 2011 Oct 21;6(1):118 [FREE Full text] [doi: [10.1186/1748-5908-6-118](https://doi.org/10.1186/1748-5908-6-118)] [Medline: [22017791](https://pubmed.ncbi.nlm.nih.gov/22017791/)]
10. Harris J, Haltbakk J, Dunning T, Austrheim G, Kirkevold M, Johnson M, et al. How patient and community involvement in diabetes research influences health outcomes: a realist review. *Health Expect* 2019 Oct 08;22(5):907-920 [FREE Full text] [doi: [10.1111/hex.12935](https://doi.org/10.1111/hex.12935)] [Medline: [31286639](https://pubmed.ncbi.nlm.nih.gov/31286639/)]
11. Lee A, Sandvei M, Asmussen HC, Skougaard M, Macdonald J, Zavada J, et al. The development of complex digital health solutions: formative evaluation combining different methodologies. *JMIR Res Protoc* 2018 Jul 16;7(7):e165 [FREE Full text] [doi: [10.2196/resprot.9521](https://doi.org/10.2196/resprot.9521)] [Medline: [30012548](https://pubmed.ncbi.nlm.nih.gov/30012548/)]
12. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012 Oct 3;21(8):1305-1314. [doi: [10.1007/s11136-011-0054-x](https://doi.org/10.1007/s11136-011-0054-x)] [Medline: [22048932](https://pubmed.ncbi.nlm.nih.gov/22048932/)]
13. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, Gibbons C, Gangannagaripalli J, Fitzpatrick R, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *J Comp Eff Res* 2016 Aug;5(5):507-519. [doi: [10.2217/cer-2015-0014](https://doi.org/10.2217/cer-2015-0014)] [Medline: [27427277](https://pubmed.ncbi.nlm.nih.gov/27427277/)]
14. Haugstvedt A, Hernar I, Graue M, Strandberg RB, Lie SS, Sigurdardottir AK, et al. Nurses' and physicians' experiences with diabetes consultations and the use of dialogue tools in the diapro pilot trial: a qualitative study. *Diabet Med* 2021 Jun 20;38(6):e14419. [doi: [10.1111/dme.14419](https://doi.org/10.1111/dme.14419)] [Medline: [33035378](https://pubmed.ncbi.nlm.nih.gov/33035378/)]

15. Skovlund PC, Ravn S, Seibaek L, Thaysen HV, Lomborg K, Nielsen BK. The development of promunication: a training-tool for clinicians using patient-reported outcomes to promote patient-centred communication in clinical cancer settings. *J Patient Rep Outcomes* 2020 Feb 11;4(1):10 [FREE Full text] [doi: [10.1186/s41687-020-0174-6](https://doi.org/10.1186/s41687-020-0174-6)] [Medline: [32048085](https://pubmed.ncbi.nlm.nih.gov/32048085/)]
16. Skovlund SE, Troelsen LH, Klim L, Jakobsen PE, Ejskjaer N. The participatory development of a national core set of person-centred diabetes outcome constructs for use in routine diabetes care across healthcare sectors. *Res Involv Engagem* 2021 Sep 10;7(1):62 [FREE Full text] [doi: [10.1186/s40900-021-00309-7](https://doi.org/10.1186/s40900-021-00309-7)] [Medline: [34507618](https://pubmed.ncbi.nlm.nih.gov/34507618/)]
17. Sheridan S, Schrandt S, Forsythe L, Hilliard TS, Paez KA, Advisory Panel on Patient Engagement (2013 inaugural panel). The PCORI engagement rubric: promising practices for partnering in research. *Ann Fam Med* 2017 Mar 13;15(2):165-170 [FREE Full text] [doi: [10.1370/afm.2042](https://doi.org/10.1370/afm.2042)] [Medline: [28289118](https://pubmed.ncbi.nlm.nih.gov/28289118/)]
18. Deane K, Delbecque L, Gorbenko O, Hamoir AM, Hoos A, Nafria B, et al. Co-creation of patient engagement quality guidance for medicines development: an international multistakeholder initiative. *BMJ Innov* 2019 Jan 02;5(1):43-55 [FREE Full text] [doi: [10.1136/bmjinnov-2018-000317](https://doi.org/10.1136/bmjinnov-2018-000317)] [Medline: [31645992](https://pubmed.ncbi.nlm.nih.gov/31645992/)]
19. Williams G, Freedman Z, Deci E. Supporting autonomy to motivate patients with diabetes for glucose control. *Diabetes Care* 1998 Oct;21(10):1644-1651. [doi: [10.2337/diacare.21.10.1644](https://doi.org/10.2337/diacare.21.10.1644)] [Medline: [9773724](https://pubmed.ncbi.nlm.nih.gov/9773724/)]
20. Kayser L, Karnoe A, Duminski E, Somekh D, Vera-Muñoz C. A new understanding of health related empowerment in the context of an active and healthy ageing. *BMC Health Serv Res* 2019 Apr 24;19(1):242 [FREE Full text] [doi: [10.1186/s12913-019-4082-5](https://doi.org/10.1186/s12913-019-4082-5)] [Medline: [31014350](https://pubmed.ncbi.nlm.nih.gov/31014350/)]
21. Gill SD, Dolley PJ, Dunning TL, Hughes AJ. Evaluating health services with point of service feedback: perspectives and experiences of patients, staff and community volunteers in an inpatient rehabilitation facility. *Disabil Rehabil* 2015 Jan 06;37(21):1997-2005. [doi: [10.3109/09638288.2014.996298](https://doi.org/10.3109/09638288.2014.996298)] [Medline: [25560371](https://pubmed.ncbi.nlm.nih.gov/25560371/)]
22. Evans C. Analysing semi-structured interviews using thematic analysis exploring voluntary civic participation among adults. *Sage Res Methods Datasets* 2018:A. [doi: [10.4135/9781526439284](https://doi.org/10.4135/9781526439284)]
23. de Visser RO, Graber R, Hart A, Abraham C, Scanlon T, Watten P, et al. Using qualitative methods within a mixed-methods approach to developing and evaluating interventions to address harmful alcohol use among young people. *Health Psychol* 2015 Apr;34(4):349-360. [doi: [10.1037/hea0000163](https://doi.org/10.1037/hea0000163)] [Medline: [25822051](https://pubmed.ncbi.nlm.nih.gov/25822051/)]
24. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. *Nurs Health Sci* 2013 Sep 11;15(3):398-405. [doi: [10.1111/nhs.12048](https://doi.org/10.1111/nhs.12048)] [Medline: [23480423](https://pubmed.ncbi.nlm.nih.gov/23480423/)]
25. Skovlund S, Nicolucci A, Balk-Møller N, Berthelsen DB, Glümer C, Perrild H, et al. Perceived benefits, barriers, and facilitators of a digital patient-reported outcomes tool for routine diabetes care: protocol for a national, multicenter, mixed methods implementation study. *JMIR Res Protoc* 2021 Sep 03;10(9):e28391 [FREE Full text] [doi: [10.2196/28391](https://doi.org/10.2196/28391)] [Medline: [34477563](https://pubmed.ncbi.nlm.nih.gov/34477563/)]
26. Mejdahl CT, Schougaard LM, Hjollund NH, Riiskjær E, Lomborg K. Patient-reported outcome measures in the interaction between patient and clinician - a multi-perspective qualitative study. *J Patient Rep Outcomes* 2020 Jan 09;4(1):3 [FREE Full text] [doi: [10.1186/s41687-019-0170-x](https://doi.org/10.1186/s41687-019-0170-x)] [Medline: [31919602](https://pubmed.ncbi.nlm.nih.gov/31919602/)]
27. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. *J Patient Rep Outcomes* 2018 Dec 15;2(1):42 [FREE Full text] [doi: [10.1186/s41687-018-0061-6](https://doi.org/10.1186/s41687-018-0061-6)] [Medline: [30294712](https://pubmed.ncbi.nlm.nih.gov/30294712/)]
28. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008 Mar 4;17(2):179-193. [doi: [10.1007/s11136-007-9295-0](https://doi.org/10.1007/s11136-007-9295-0)] [Medline: [18175207](https://pubmed.ncbi.nlm.nih.gov/18175207/)]
29. Holmes MM, Lewith G, Newell D, Field J, Bishop FL. The impact of patient-reported outcome measures in clinical practice for pain: a systematic review. *Qual Life Res* 2017 Feb 4;26(2):245-257 [FREE Full text] [doi: [10.1007/s11136-016-1449-5](https://doi.org/10.1007/s11136-016-1449-5)] [Medline: [27815820](https://pubmed.ncbi.nlm.nih.gov/27815820/)]
30. Grimaldi A, Penforinis A, Consoli S, Falissard B, Eymard E, Williams P, et al. Breaking barriers to effective type 2 diabetes management: findings from the use of the optima© questionnaire in clinical practice. *Adv Ther* 2016 Jun 18;33(6):1033-1048 [FREE Full text] [doi: [10.1007/s12325-016-0341-6](https://doi.org/10.1007/s12325-016-0341-6)] [Medline: [27193870](https://pubmed.ncbi.nlm.nih.gov/27193870/)]
31. Rodriguez H, Glenn B, Olmos T, Krist A, Shimada S, Kessler R. Real-world implementation and outcomes of health behavior and mental health assessment. *J Am Board Fam Med Internet*. 2014. URL: <http://www.jabfm.org/cgi/doi/10.3122/jabfm.2014.03.130264> [accessed 2019-03-25]
32. Eilander M, de Wit M, Rotteveel J, van Schaaik NM, Roeleveld-Versteegh A, Snoek F. Implementation of quality of life monitoring in dutch routine care of adolescents with type 1 diabetes: appreciated but difficult. *Pediatr Diabetes* 2016 Mar 07;17(2):112-119. [doi: [10.1111/pedi.12237](https://doi.org/10.1111/pedi.12237)] [Medline: [25580639](https://pubmed.ncbi.nlm.nih.gov/25580639/)]
33. Scholle S, Morton S, Homco J, Rodriguez K, Anderson D, Hahn E, et al. Implementation of the PROMIS-29 in routine care for people with diabetes. *J Ambul Care Manage* 2018 Oct;41(4):274-287. [doi: [10.1097/jac.0000000000000248](https://doi.org/10.1097/jac.0000000000000248)]
34. Zhang J, Chen Y, Ashfaq S, Bell K, Calvitti A, Farber N, et al. Strategizing EHR use to achieve patient-centered care in exam rooms: a qualitative study on primary care providers. *J Am Med Inform Assoc* 2016 Jan;23(1):137-143 [FREE Full text] [doi: [10.1093/jamia/ocv142](https://doi.org/10.1093/jamia/ocv142)] [Medline: [26568605](https://pubmed.ncbi.nlm.nih.gov/26568605/)]
35. Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'as' for patient-centered counseling. *Health Promot Int* 2006 Sep 04;21(3):245-255. [doi: [10.1093/heapro/dal017](https://doi.org/10.1093/heapro/dal017)] [Medline: [16751630](https://pubmed.ncbi.nlm.nih.gov/16751630/)]

36. Dickinson WP, Dickinson LM, Jortberg BT, Hessler DM, Fernald DH, Fisher L. A protocol for a cluster randomized trial comparing strategies for translating self-management support into primary care practices. *BMC Fam Pract* 2018 Jul 24;19(1):126 [FREE Full text] [doi: [10.1186/s12875-018-0810-x](https://doi.org/10.1186/s12875-018-0810-x)] [Medline: [30041598](https://pubmed.ncbi.nlm.nih.gov/30041598/)]
37. Larivaara P, Kiuttu J, Taanila A. The patient-centred interview: the key to biopsychosocial diagnosis and treatment. *Scand J Prim Health Care* 2001 Mar;19(1):8-13. [doi: [10.1080/028134301300034521](https://doi.org/10.1080/028134301300034521)] [Medline: [11303552](https://pubmed.ncbi.nlm.nih.gov/11303552/)]
38. Dickinson JK, Guzman SJ, Maryniuk MD, O'Brian CA, Kadohiro JK, Jackson RA, et al. The use of language in diabetes care and education. *Diabetes Educ* 2017 Dec 17;43(6):551-564. [doi: [10.1177/0145721717735535](https://doi.org/10.1177/0145721717735535)] [Medline: [29040034](https://pubmed.ncbi.nlm.nih.gov/29040034/)]
39. Sturt J, Dennick K, Due-Christensen M, McCarthy K. The detection and management of diabetes distress in people with type 1 diabetes. *Curr Diab Rep* 2015 Nov 28;15(11):101. [doi: [10.1007/s11892-015-0660-z](https://doi.org/10.1007/s11892-015-0660-z)] [Medline: [26411924](https://pubmed.ncbi.nlm.nih.gov/26411924/)]
40. Reddy J, Wilhelm K, Campbell L. Putting paid to diabetes-related distress: the potential utility of the problem areas in diabetes (PAID) scale in patients with diabetes. *Psychosomatics* 2013 Jan;54(1):44-51. [doi: [10.1016/j.psym.2012.08.004](https://doi.org/10.1016/j.psym.2012.08.004)] [Medline: [23295006](https://pubmed.ncbi.nlm.nih.gov/23295006/)]
41. Frost MH, Bonomi AE, Cappelleri JC, Schünemann HJ, Moynihan TJ, Aaronson NK, Clinical Significance Consensus Meeting Group. Applying quality-of-life data formally and systematically into clinical practice. *Mayo Clin Proc* 2007 Oct;82(10):1214-1228 [FREE Full text] [doi: [10.4065/82.10.1214](https://doi.org/10.4065/82.10.1214)] [Medline: [17908528](https://pubmed.ncbi.nlm.nih.gov/17908528/)]
42. Pedersen-Bjergaard U, Agerholm-Larsen B, Pramming S, Hougaard P, Thorsteinsson B. Activity of angiotensin-converting enzyme and risk of severe hypoglycaemia in type 1 diabetes mellitus. *Lancet* 2001 Apr;357(9264):1248-1253. [doi: [10.1016/s0140-6736\(00\)04405-6](https://doi.org/10.1016/s0140-6736(00)04405-6)]
43. Pedersen-Bjergaard U, Thorsteinsson B. Reporting severe hypoglycemia in type 1 diabetes: facts and pitfalls. *Curr Diab Rep* 2017 Oct 28;17(12):131. [doi: [10.1007/s11892-017-0965-1](https://doi.org/10.1007/s11892-017-0965-1)] [Medline: [29080929](https://pubmed.ncbi.nlm.nih.gov/29080929/)]
44. Arbuckle RA, Humphrey L, Vardeva K, Arondekar B, Danten-Viala M, Scott JA, et al. Psychometric evaluation of the diabetes symptom checklist-revised (DSC-R)--a measure of symptom distress. *Value Health* 2009 Nov;12(8):1168-1175 [FREE Full text] [doi: [10.1111/j.1524-4733.2009.00571.x](https://doi.org/10.1111/j.1524-4733.2009.00571.x)] [Medline: [19558371](https://pubmed.ncbi.nlm.nih.gov/19558371/)]
45. Hajos TR, Pouwer F, Skovlund SE, Den Oudsten BL, Geelhoed-Duijvestijn PH, Tack CJ, et al. Psychometric and screening properties of the WHO-5 well-being index in adult outpatients with type 1 or type 2 diabetes mellitus. *Diabet Med* 2013 Feb 21;30(2):63-69. [doi: [10.1111/dme.12040](https://doi.org/10.1111/dme.12040)] [Medline: [23072401](https://pubmed.ncbi.nlm.nih.gov/23072401/)]
46. Foot H, La Caze A, Gujral G, Cottrell N. The necessity-concerns framework predicts adherence to medication in multiple illness conditions: a meta-analysis. *Patient Educ Couns* 2016 May;99(5):706-717. [doi: [10.1016/j.pec.2015.11.004](https://doi.org/10.1016/j.pec.2015.11.004)] [Medline: [26613666](https://pubmed.ncbi.nlm.nih.gov/26613666/)]
47. Horne R, Chapman SC, Parham R, Freemantle N, Forbes A, Cooper V. Understanding patients' adherence-related beliefs about medicines prescribed for long-term conditions: a meta-analytic review of the necessity-concerns framework. *PLoS One* 2013 Dec 2;8(12):e80633 [FREE Full text] [doi: [10.1371/journal.pone.0080633](https://doi.org/10.1371/journal.pone.0080633)] [Medline: [24312488](https://pubmed.ncbi.nlm.nih.gov/24312488/)]
48. Connor U, Kessler L, de Groot M, Neill R, Sandy R. Implementing person-centered communication in diabetes care: a new tool for diabetes care professionals. *Patient Prefer Adherence* 2019 Aug;Volume 13:1443-1450. [doi: [10.2147/ppa.s214092](https://doi.org/10.2147/ppa.s214092)]
49. Grimaldi A, Penforis A, Consoli S, Falissard B, Eymard E, Williams P, et al. Breaking barriers to effective type 2 diabetes management: findings from the use of the OPTIMA© questionnaire in clinical practice. *Adv Ther* 2016 Jun;33(6):1033-1048 [FREE Full text] [doi: [10.1007/s12325-016-0341-6](https://doi.org/10.1007/s12325-016-0341-6)] [Medline: [27193870](https://pubmed.ncbi.nlm.nih.gov/27193870/)]
50. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, et al. Management of hyperglycaemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetologia* 2018 Dec 5;61(12):2461-2498. [doi: [10.1007/s00125-018-4729-5](https://doi.org/10.1007/s00125-018-4729-5)] [Medline: [30288571](https://pubmed.ncbi.nlm.nih.gov/30288571/)]
51. Jones A, Vallis M, Pouwer F. If it does not significantly change HbA1c levels why should we waste time on it? A plea for the prioritization of psychological well-being in people with diabetes. *Diabet Med* 2015 Feb 10;32(2):155-163. [doi: [10.1111/dme.12620](https://doi.org/10.1111/dme.12620)] [Medline: [25354315](https://pubmed.ncbi.nlm.nih.gov/25354315/)]
52. Bajracharya P, Summers L, Amatya AK, DeBlicek C. Implementation of a depression screening protocol and tools to improve screening for depression in patients with diabetes in the primary care setting. *J Nurse Pract* 2016 Nov;12(10):690-696. [doi: [10.1016/j.nurpra.2016.08.009](https://doi.org/10.1016/j.nurpra.2016.08.009)]
53. Dennick K, Sturt J, Speight J. What is diabetes distress and how can we measure it? A narrative review and conceptual model. *J Diabetes Complications* 2017 May;31(5):898-911. [doi: [10.1016/j.jdiacomp.2016.12.018](https://doi.org/10.1016/j.jdiacomp.2016.12.018)] [Medline: [28274681](https://pubmed.ncbi.nlm.nih.gov/28274681/)]
54. Peyrot M, Burns KK, Davies M, Forbes A, Hermanns N, Holt R, et al. Diabetes Attitudes Wishes and Needs 2 (DAWN2): a multinational, multi-stakeholder study of psychosocial issues in diabetes and person-centred diabetes care. *Diabetes Res Clin Pract* 2013 Feb;99(2):174-184 [FREE Full text] [doi: [10.1016/j.diabres.2012.11.016](https://doi.org/10.1016/j.diabres.2012.11.016)] [Medline: [23273515](https://pubmed.ncbi.nlm.nih.gov/23273515/)]
55. Glasgow RE, Nelson CC, Strycker LA, King DK. Using re-aim metrics to evaluate diabetes self-management support interventions. *Am J Prev Med* 2006 Jan;30(1):67-73. [doi: [10.1016/j.amepre.2005.08.037](https://doi.org/10.1016/j.amepre.2005.08.037)] [Medline: [16414426](https://pubmed.ncbi.nlm.nih.gov/16414426/)]
56. Krist AH, Phillips SM, Sabo RT, Balasubramanian BA, Heurtin-Roberts S, Ory MG, MOHR Study Group. Adoption, reach, implementation, and maintenance of a behavioral and mental health assessment in primary care. *Ann Fam Med* 2014 Nov 10;12(6):525-533 [FREE Full text] [doi: [10.1370/afm.1710](https://doi.org/10.1370/afm.1710)] [Medline: [25384814](https://pubmed.ncbi.nlm.nih.gov/25384814/)]
57. Forman J, Heisler M, Damschroder LJ, Kaselitz E, Kerr EA. Development and application of the re-aim quest mixed methods framework for program evaluation. *Prev Med Rep* 2017 Jun;6:322-328 [FREE Full text] [doi: [10.1016/j.pmedr.2017.04.002](https://doi.org/10.1016/j.pmedr.2017.04.002)] [Medline: [28451518](https://pubmed.ncbi.nlm.nih.gov/28451518/)]

58. Scholle S, Morton S, Homco J, Rodriguez K, Anderson D, Hahn E. Implementation of the PROMIS-29 in routine care for people with diabetes challenges and opportunities. *J Ambul Care Manage* 2018;41(4):87. [doi: [10.1097/jac.0000000000000248](https://doi.org/10.1097/jac.0000000000000248)]
59. Eilander M, de Wit M, Rotteveel J, van Schaaik NM, Roeleveld-Versteegh A, Snoek F. Implementation of quality of life monitoring in Dutch routine care of adolescents with type 1 diabetes: appreciated but difficult. *Pediatr Diabetes* 2016 Mar 07;17(2):112-119. [doi: [10.1111/vedi.12237](https://doi.org/10.1111/vedi.12237)] [Medline: [25580639](https://pubmed.ncbi.nlm.nih.gov/25580639/)]
60. Wieringa TH, Rodriguez-Gutierrez R, Spencer-Bonilla G, de Wit M, Ponce OJ, Sanchez-Herrera MF, et al. Decision aids that facilitate elements of shared decision making in chronic illnesses: a systematic review. *Syst Rev* 2019 May 20;8(1):121 [FREE Full text] [doi: [10.1186/s13643-019-1034-4](https://doi.org/10.1186/s13643-019-1034-4)] [Medline: [31109357](https://pubmed.ncbi.nlm.nih.gov/31109357/)]
61. Ye Q, Patel R, Khan U, Boren SA, Kim MS. Evaluation of provider documentation patterns as a tool to deliver ongoing patient-centred diabetes education and support. *Int J Clin Pract* 2020 Mar 21;74(3):e13451 [FREE Full text] [doi: [10.1111/ijcp.13451](https://doi.org/10.1111/ijcp.13451)] [Medline: [31769903](https://pubmed.ncbi.nlm.nih.gov/31769903/)]
62. Kaplan SH, Billimek J, Sorkin DH, Ngo-Metzger Q, Greenfield S. Reducing racial/ethnic disparities in diabetes: the coached care (R2D2C2) project. *J Gen Intern Med* 2013 Oct 4;28(10):1340-1349 [FREE Full text] [doi: [10.1007/s11606-013-2452-y](https://doi.org/10.1007/s11606-013-2452-y)] [Medline: [23645452](https://pubmed.ncbi.nlm.nih.gov/23645452/)]
63. Ritter PL, Lorig K, Laurent DD. Characteristics of the Spanish- and English-language self-efficacy to manage diabetes scales. *Diabetes Educ* 2016 Apr 04;42(2):167-177. [doi: [10.1177/0145721716628648](https://doi.org/10.1177/0145721716628648)] [Medline: [26846336](https://pubmed.ncbi.nlm.nih.gov/26846336/)]
64. Anderson RM, Funnell MM, Fitzgerald JT, Marrero DG. The diabetes empowerment scale: a measure of psychosocial self-efficacy. *Diabetes Care* 2000 Jun 01;23(6):739-743 [FREE Full text] [doi: [10.2337/diacare.23.6.739](https://doi.org/10.2337/diacare.23.6.739)] [Medline: [10840988](https://pubmed.ncbi.nlm.nih.gov/10840988/)]
65. American Association of Diabetes Educators. An effective model of diabetes care and education: revising the AADE7 Self-Care Behaviors®. *Diabetes Educ* 2020 Apr 12;46(2):139-160. [doi: [10.1177/0145721719894903](https://doi.org/10.1177/0145721719894903)] [Medline: [31928334](https://pubmed.ncbi.nlm.nih.gov/31928334/)]
66. Shillington A, Ganjuli A, Clewell J. The impact of patient support programs on adherence, clinical, humanistic, and economic patient outcomes: a targeted systematic review. *Patient Prefer Adherence* 2016 Apr:711. [doi: [10.2147/ppa.s101175](https://doi.org/10.2147/ppa.s101175)]
67. Glasgow R, Whitesides H, Nelson C, King D. Use of the patient assessment of chronic illness care (PACIC) with diabetic patients: relationship to patient characteristics, receipt of care, and self-management. *Diabetes Care*. 2005. URL: <http://care.diabetesjournals.org/cgi/doi/10.2337/diacare.28.11.2655> [accessed 2019-03-25]
68. Brownson CA, Miller D, Crespo R, Neuner S, Thompson J, Wall JC, et al. A quality improvement tool to assess self-management support in primary care. *Jt Comm J Qual Patient Saf* 2007 Jul;33(7):408-416. [doi: [10.1016/s1553-7250\(07\)33047-x](https://doi.org/10.1016/s1553-7250(07)33047-x)] [Medline: [17711143](https://pubmed.ncbi.nlm.nih.gov/17711143/)]
69. Nuño R, Coleman K, Bengoa R, Sauto R. Integrated care for chronic conditions: the contribution of the ICCC framework. *Health Policy* 2012 Apr;105(1):55-64. [doi: [10.1016/j.healthpol.2011.10.006](https://doi.org/10.1016/j.healthpol.2011.10.006)] [Medline: [22071454](https://pubmed.ncbi.nlm.nih.gov/22071454/)]
70. Porter M. Value-based health care delivery : outcomes measurement creating a value-based health care delivery organization the strategic agenda. *Ann Surg* 2008;248(4):9. [doi: [10.1097/sla.0b013e31818a43af](https://doi.org/10.1097/sla.0b013e31818a43af)]
71. Paradis E, Sutkin G. Beyond a good story: from hawthorne effect to reactivity in health professions education research. *Med Educ* 2017 Jan 31;51(1):31-39. [doi: [10.1111/medu.13122](https://doi.org/10.1111/medu.13122)] [Medline: [27580703](https://pubmed.ncbi.nlm.nih.gov/27580703/)]

Abbreviations

HCP: health care professional

PRO: patient-reported outcome

RE-AIM: reach, efficacy, adoption, implementation, and maintenance

VBHC-PRO-DIA: Value-Based Health Care and Patient-Reported Outcome in Diabetes

Edited by G Eysenbach; submitted 02.03.21; peer-reviewed by D Chao, K Smith; comments to author 15.05.21; revised version received 19.06.21; accepted 24.08.21; published 03.11.21.

Please cite as:

Skovlund SE, Troelsen LH, Noergaard LM, Pietraszek A, Jakobsen PE, Ejksjaer N

Feasibility and Acceptability of a Digital Patient-Reported Outcome Tool in Routine Outpatient Diabetes Care: Mixed Methods Formative Pilot Study

JMIR Form Res 2021;5(11):e28329

URL: <https://formative.jmir.org/2021/11/e28329>

doi: [10.2196/28329](https://doi.org/10.2196/28329)

PMID: [34730545](https://pubmed.ncbi.nlm.nih.gov/34730545/)

©Soren E Skovlund, Lise Havbæk Troelsen, Lise Møllergaard Noergaard, Anna Pietraszek, Poul Erik Jakobsen, Niels Ejksjaer. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.11.2021. 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

Usability Evaluation of an Offline Electronic Data Capture App in a Prospective Multicenter Dementia Registry (digiDEM Bayern): Mixed Method Study

Michael Reichold¹, MSc; Miriam Heß¹, MSc; Peter Kolominsky-Rabas², Prof Dr; Elmar Gräbel³, Prof Dr; Hans-Ulrich Prokosch¹, Prof Dr

¹Department of Medical Informatics, Biometrics and Epidemiology, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

²Interdisciplinary Center for Health Technology Assessment and Public Health (IZPH), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

³Center for Health Services Research in Medicine, Department of Psychiatry and Psychotherapy, University Hospital Erlangen, Erlangen, Germany

Corresponding Author:

Michael Reichold, MSc

Department of Medical Informatics, Biometrics and Epidemiology

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Wetterkreuz 15

Erlangen, 91058

Germany

Phone: 49 91318526720

Email: michael.reichold@fau.de

Abstract

Background: Digital registries have been shown to provide an efficient way of gaining a better understanding of the clinical complexity and long-term progression of diseases. The paperless method of electronic data capture (EDC) during a patient interview saves both time and resources. In the prospective multicenter project “Digital Dementia Registry Bavaria (digiDEM Bayern),” interviews are also performed on site in rural areas with unreliable internet connectivity. It must be ensured that EDC can still be performed in such a context and that there is no need to fall back on paper-based questionnaires. In addition to a web-based data collection solution, the EDC system REDCap (Research Electronic Data Capture) offers the option to collect data offline via an app and to synchronize it afterward.

Objective: The aim of this study was to evaluate the usability of the REDCap app as an offline EDC option for a lay user group and to examine the necessary technology acceptance of using mobile devices for data collection. The feasibility of the app-based offline data collection in the digiDEM Bayern dementia registry project was then evaluated before going live.

Methods: An exploratory mixed method design was employed in the form of an on-site usability test with the “Thinking Aloud” method combined with an online questionnaire including the System Usability Scale (SUS). The acceptance of mobile devices for data collection was surveyed based on five categories of the technology acceptance model.

Results: Using the “Thinking Aloud” method, usability issues were identified and solutions were accordingly derived. Evaluation of the REDCap app resulted in a SUS score of 74, which represents “good” usability. After evaluating the technology acceptance questionnaire, it can be concluded that the lay user group is open to mobile devices as interview tools.

Conclusions: The usability evaluation results show that a lay user group generally agree that data collecting partners in the digiDEM project can handle the REDCap app well. The usability evaluation provided statements about positive aspects and could also identify usability issues relating to the REDCap app. In addition, the current technology acceptance in the sample showed that heterogeneous groups of different ages with diverse experiences in handling mobile devices are also ready for the use of app-based EDC systems. Based on these results, it can be assumed that the offline use of an app-based EDC system on mobile devices is a viable solution for collecting data in a decentralized registry-based research project.

(JMIR Form Res 2021;5(11):e31649) doi:[10.2196/31649](https://doi.org/10.2196/31649)

KEYWORDS

dementia; usability; evaluation; mobile device; registry; electronic data collection; offline; mobile app; digital health; usability testing

Introduction

Patient registries have proven to be valuable tools for public health surveillance and research studies [1]. In these organized databases, observational study methods are used to collect uniform data and to evaluate specified outcomes for a defined population [2]. As an alternative to paper-based data collection, electronic data capture (EDC) systems have become established in registries in recent years [3]. In particular, EDC systems with web-based data collection tools have been widely accepted as the state-of-the-art approach in multicenter studies [4,5]. Such systems offer several advantages, including time and cost savings [6,7] or higher data quality [8-10] such as through plausibility checks during data entry. However, to make the most of these benefits, the EDC system must be designed for usability and tailored to the user's preferred method of data collection [11].

To successfully integrate an EDC system into a registry-based research project, it is essential to coordinate the system requirements and usability with future users in advance [12,13]. Therefore, the usability of a system is critical to the success of interactive applications in health care [11]. Only a system that users see as fit for purpose has a better chance of being accepted and used in the long term [14,15]. In the fields of eHealth [16,17] and mobile health [18-20], user-centered usability studies are widely used to identify problems in systems. Although Schmier et al [21] already demonstrated the practical applications of usability theory for EDC systems in clinical trials in 2005, usability studies in EDC-based registry research are still comparatively rare [22].

The project "Digital Dementia Registry Bavaria (digiDEM Bayern)" [23] aims to establish a multicenter, prospective, longitudinal dementia registry. The data collection process is organized with a decentralized approach, including staff from different types of outpatient counseling and caring institutions distributed across Bavaria (Germany). Data collection often takes place in rural areas in the homes of people with dementia, where internet access is patchy or unavailable. In these cases, the project will gather electronic data offline using mobile devices and the app of the EDC system known as REDCap (Research Electronic Data Capture) [24,25].

Thus, the objectives of this study were to evaluate the usability of the REDCap app as an option for offline electronic data collection and to examine whether the target user group has the necessary technology acceptance for data collection using mobile devices. This study should help to identify potential barriers and evaluate the feasibility of the REDCap app in a registry study with a large number of distributed nonexpert data collection partners and the need for offline on-site data collection.

Methods

Setting

To foster dementia research, the registry project digiDEM Bayern [23] collects data from people with mild cognitive impairment or mild-to-moderate dementia and their family caregivers over a period of 3 years throughout all seven administrative districts of Bavaria. The findings will help to improve the living conditions of people with dementia and their caregiving relatives, especially in rural areas of Bavaria.

In the digiDEM project, data collection for the registry is carried out by approximately 300 so-called "digiDEM partners," who are employees (such as nursing assistants and home health aides) from, for example, community counseling bodies, memory clinics, daycare facilities, or outpatient care organizations distributed across Bavaria that counsel or care for people with dementia and family caregivers. During face-to-face surveys involving interviews of people with dementia and their caregivers, the digiDEM partners enter various types of information [23], including data about diagnosis, cognitive trajectories, behavioral and psychological symptoms, and the care situation, into a web-based EDC system (REDCap). However, owing to the lack of mobility or poor health of people with dementia [26,27], conducting the survey at the digiDEM partner's facility is not always possible. In these cases, the data must be gathered in the homes of people with dementia. Given the fact that access to the internet in Germany decreases with age (64.4% at age 73-78 years, 39.4% at age 79-84 years) and when people are living in rural areas [28,29], there is no guarantee that the digiDEM partners can use an existing on-site internet connection for the web-based EDC system while undertaking the survey at participants' homes.

System Description

In digiDEM, we use REDCap as our EDC system [24]. REDCap is a secure, web-based software platform designed to support data capture for research studies. In addition to the direct web-based data collection, REDCap offers the option of collecting data offline via the REDCap mobile app. The data can be synchronized to the central registry database subsequently [30]. REDCap has been adopted by more than 5265 partners in 142 countries since its initial development at Vanderbilt University [31].

In our usability evaluation, the app was used "out of the box" as offered by REDCap. The rendering style of the app was retained at the default setting "New render form." Other than the current German language pack "German (DE)" being activated, no further adjustments were made to the REDCap app. The app is provided in English by default; however, there is an option to activate a language file to translate the user interface. Unfortunately, the German language files do not yet cover a full user interface translation (for some screenshots, see [Multimedia Appendix 1](#)). Therefore, some system messages

are still in English, such as the synchronization report or error messages caused by missing values for mandatory fields. The questionnaire and project’s own customized warning messages for plausibility checks can be created in the REDCap designer function and are displayed in German.

Material

Participants were provided with a tablet (Apple iPad Air 2 with iOS 13.7) to carry out the predefined tasks. The REDCap app (version 5.9.6) was preinstalled on the tablet and a dummy registry project with the test questionnaire for the usability study was set up. A user account was set up in advance for the participants. The test questionnaire contains a subset of questions from the original digiDEM questionnaire [23]. When designing the questionnaire, care was taken to use all field types that had also been used in the digiDEM questionnaire (textbox, drop-down list, radio buttons, multiple-choice checkboxes, date field). Some questions were linked by means of branching logic only to appear if the previous question was answered with a defined value. A plausibility warning message was also included to monitor the participants’ reactions to such warnings.

Study Design

We performed a mixed methods study in our usability evaluation with an exploratory sequential design [32]. The qualitatively driven study (known as “Quant->qual”) started with a “Thinking Aloud” [33] component, followed by the standardized quantitative System Usability Scale (SUS) usability questionnaire. To also examine participants’ attitudes toward the mobile device–based data collection, a questionnaire based on the technology acceptance model (TAM) was applied. The study was performed over 4 weeks in October 2020.

Using a mixed methods approach, we were able to triangulate several data sources and hence consider the research question from different perspectives [34]: participants’ thoughts, demographic data, and structured questionnaires. Combining data from qualitative and quantitative methods can balance the strengths and offset possible limitations of the respective method [17], provide a more comprehensive understanding of evidence [35], and help researchers investigate the usability as comprehensively as possible [36,37].

Test Procedure

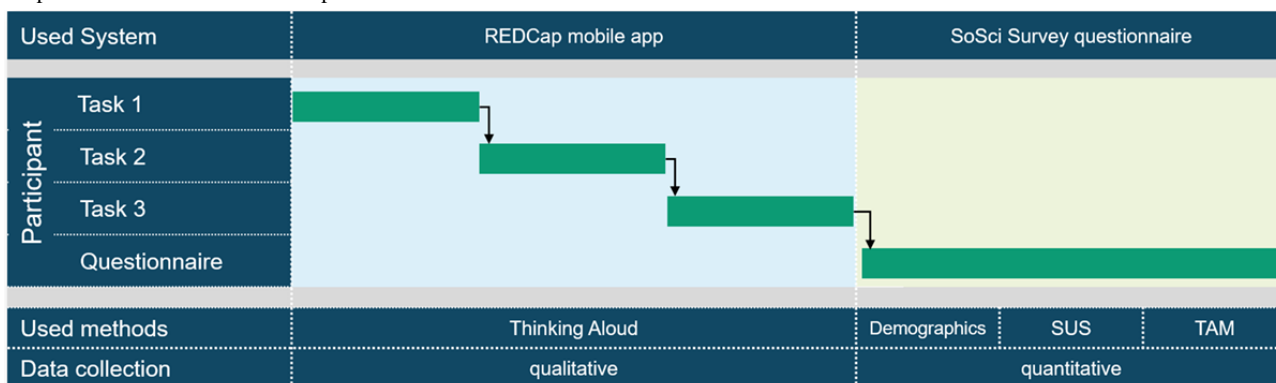
Design

The test procedure consisted of two sequential independent parts, as illustrated in Figure 1. In the first part, qualitative data were collected using the Thinking Aloud method. In the second part, quantitative data were obtained using an online questionnaire.

We pretested the test procedure with four scientific project members to determine its suitability for obtaining rich data to address the proposed research objectives. Furthermore, technical and operational problems were addressed so that they could be excluded during the test.

Before starting the test procedure, the participants were informed about the app’s purpose and the test procedure. A demonstration video was produced in advance and shown before the test to familiarize participants with the Thinking Aloud method. In addition, a test manual was created, including all questions (for the participant) and answers (for the simulated patient) for the tasks in the test survey (Multimedia Appendix 2). To ensure that participants had sufficient information to perform the tasks, they received a brief in-person tutorial on using the REDCap app (Multimedia Appendix 3).

Figure 1. Overview of the systems and methods used during the test procedure. SUS: System Usability Scale; TAM: technology acceptance model; REDCap: Research Electronic Data Capture.



Qualitative Data Collection

Quantitative usability questionnaires cannot provide precise information about why a participant has rated usability in a particular way, and no direct usability problems can be derived from these responses [37]. Therefore, we used the Thinking Aloud method as a user-centered design approach to derive qualitative statements [38]. There is a potential risk of the participant forgetting to simultaneously express their thoughts while solving a task [39]. If a researcher actively intervenes and

asks them to explain their thoughts, this can distract the participant’s attention or modify their thought processes. This interference can be seen as a kind of experimenter bias, which can strongly impact the reliability and validity of the qualitative data [40,41]. Therefore, the research team ensured as little interference as possible within the Thinking Aloud process [11]. Participants were simply reminded to keep talking if they stopped verbalizing their thoughts. Other, more intrusive types of probes to gather even more useful information were not used.

For the Thinking Aloud test, a digiDEM on-site interview situation was simulated. The participant had to enter the data into the app while interviewing people with dementia. The interviewee was simulated by a research assistant and was the same person for all participants. In the test survey, the participant had to complete three predefined tasks in the REDCap app. The tasks increased in complexity and represented realistic examples of tasks in the digiDEM data collection pathway. The first task required offline data collection in the form of a baseline interview, which included questions such as “What is your marital status?” or “Is there a medically confirmed diagnosis of dementia?” (Multimedia Appendix 1). The participant then had to upload the offline collected data to the server (synchronization) in the second task. In the final task, the participant conducted a follow-up interview.

Quantitative Data Collection

After completing the three tasks, the participant had to fill out an online questionnaire by means of the SoSci Survey tool [42], which consisted of sociodemographic data, the SUS [43], and a detailed TAM questionnaire (Multimedia Appendix 4).

The sociodemographic part included three closed questions on age, gender, and experience with technical devices. The SUS is a standardized scoring questionnaire that ensures a valid and reliable measurement for usability [44,45]. It provides usable results even with a smaller sample [46]. The German version of the SUS questionnaire was used in this study [47].

To rule out the possibility that a dismissive attitude toward mobile devices for data collection leads to poor usability, we evaluated technology acceptance based on the TAM. According to Davis [48], two factors are crucial for determining technology acceptance: perceived usefulness and perceived ease of use. Although the TAM is a well-established instrument for determining technology acceptance in health care [49], according to Holden and Karsh [50] and Ammenwerth [51], the model could benefit from modifications by taking into account external influencing variables in the health care environment. To account for this possibility, we included three further categories. The category “anxiety” was added [52] since some users experience anxiety when they are asked to use a new system [53]. The second category added was “social influence” [54]; this was added because digiDEM partners often work together in teams at their institution, and therefore the social environment could have an impact on acceptance [55,56]. Third, since the successful use of technology depends heavily on adequate organizational and technical infrastructure and support, the category “facilitating conditions” was also included [57]. Therefore, our TAM questionnaire contained five categories (perceived usefulness, perceived ease of use, anxiety, social influence, and facilitating conditions), each evaluated on a 5-point Likert scale. Each category was investigated by way of multiple items (Multimedia Appendix 4).

Participants and Recruitment

We took care that the sociodemographic characteristics of the sample were as broadly distributed as possible to guarantee validity and trustworthiness of the qualitative data [58,59]. In particular, all age groups needed to be included in the sample,

as age can significantly affect the usability rating [60,61]. For the sample selection, the digiDEM partners were first divided into subgroups based on the type of facility in which they worked: community counseling bodies, support groups, flat-sharing communities, daycare facilities, outpatient care organizations, geriatric rehabilitation facilities, and research institutes. From each subgroup, 1-2 participants were recruited randomly using balanced randomization.

We estimated requiring a minimum sample size of 12 participants based on our test procedure for a proper usability evaluation. Because the Thinking Aloud method provides a rich source of qualitative data [11], even a small sample (approximately 8 participants) is sufficient to understand the task behavior [62] and to identify the main usability problems [63]. For quantitative data, Tullis and Stetson [46] observed that even with 12 participants, the SUS questionnaire produced the same results as a larger sample in at least 90% of the cases studied.

Inclusion and Exclusion Criteria

Participants in the study were selected from digiDEM partners who will eventually carry out the data collection as part of the digiDEM project. There were no limitations based on age, profession, or experience with information technology. Because digiDEM partners who have already gained experience with the REDCap app or another EDC system had to be excluded, we did not include memory clinic facilities that had participated in an earlier study.

Data Analysis

Qualitative Data Analysis

The qualitative evaluation of the Thinking Aloud results was based on content analysis. Therefore, the participants were filmed while performing the tasks and a screen capture video of the tablet was recorded. All recordings of the Thinking Aloud test were transcribed verbatim and analyzed according to the structured content analysis method developed by Mayring [64]. The software MAXQDA (Version 2020 Plus) was used to transcribe the recordings and analyze data.

To ensure the trustworthiness of the qualitative data, we followed the checklist drawn up by Elo et al [65]. Accordingly, we started with a preliminary content analysis after the first participants had completed the test procedure. Two researchers (MR, MH) independently coded four interviews. To increase reliability, coders differed in age, gender, and professional background [64,66]. The participants’ statements were divided into the following main categories: “positive aspects,” “suggestions for improvement,” and “problems.” In line with the Zapf taxonomy of errors, the category “problems” was subdivided into either “functionality problems” or “usability problems” [67]. Functionality problems refer to the mismatch between the task being carried out and the app (eg, an error occurring while data are being uploaded), whereas usability problems refer to the mismatch between the user and app (eg, the app does not fit the user’s expectation because the logout button is only found on the landing page) [68,69].

Subcategories for the main categories were defined during the analysis. This category system served as the basis for coding the remaining transcripts. Given the descriptive nature of the data, additional subcategories emerged during coding. This process continued until saturation of the category system was achieved [58]. In a second run, a complete back check of the designed structure was performed. Finally, to increase trustworthiness, the entire research team reviewed the analysis process and categorizations in terms of researcher triangulation [70]. Differences in the coding were discussed and resolved by mutual agreement [71].

Because issues in the category “usability problems” mainly influence software usability, these statements were weighted by two independent researchers (MR, MH) according to the severity rating formulated by Nielsen [62,63]. The severity rating scale ranges from 0 to 4, where 0 means “I don’t agree that this is a usability problem at all” and 4 means “Usability catastrophe: imperative to fix this before product can be released” [62]. Major issues should be given higher priority because they can impact data quality, satisfaction, and functionality [22]. As Nielsen did not specify a method for calculating an overall severity score, we calculated an issue’s severity score by multiplying the severity rating by the number of mentions.

Quantitative Data Analysis

The SUS was evaluated using Brooke’s evaluation scheme [43]. Therefore, the weighting of the answers (1=strongly disagree to 5=strongly agree) was recoded. For positively worded questions, the most negative answer option (strongly disagree) was weighted with a 0 and the most positive answer (strongly agree) was weighted with a 4. For negatively worded questions, the coding was the exact reverse. All values per subject were summed and multiplied by 2.5. Thus, a value ranging from 0 (worst imaginable usability) to 100 (best imaginable usability) was achieved, ensuring the values’ comparability [43]. The

boxplot form was chosen to present the data distribution based on the results of the TAM questionnaire in a standardized manner. All statistical calculations were performed with SPSS (Version 26.0) software.

Ethical Statement

Before this study, approval was obtained from the institutional review board, the Committee on Research Ethics of the Friedrich-Alexander-University Erlangen-Nürnberg (Germany), following all applicable regulations (346_20 Bc). Informed consent was obtained in writing from all participants beforehand. Participation in the study was voluntary and no incentives were offered for participating.

Results

Participant Characteristics

In total, 12 participants took part in our usability study (6 men and 6 women). The participants were distributed over different types of institutions: community counseling (2), support groups (1), flat-sharing communities (1), daycare facilities (2), outpatient care organizations (2), geriatric rehabilitation facilities (2), and research institutes (2). The age of the participants covered all five age groups: 18-24 years, 25-34 years, 35-44 years, 45-49 years, and >60 years (Table 1).

When asked about experience with mobile devices (personal or professional), 10 out of 12 participants mentioned experience with a smartphone, half of the sample had experience using tablets, and one participant had no experience with any of the listed devices (Table 1).

None of the participants had any experience in EDC systems or registry-based research studies. These were medical assistants, nursing assistants, home health aides, and volunteer assistants, with caring for or counseling people with dementia and family caregivers as their primary role.

Table 1. Characteristics of the participants: age group, gender, and device experience.

Characteristics	Women (n=6), n	Men (n=6), n	Total (N=12), n
Age group (years)			
18-24	1	0	1
25-34	2	2	4
35-44	0	1	1
45-59	2	3	5
>60	1	0	1
Device experience (private or professional)			
Smartphone	6	4	10
Tablet	2	4	6
Desktop-PC	4	4	8
Laptop	6	3	9
None	0	1	1

Thinking Aloud Test

Overview

The time it took participants to complete the three tasks in the REDCap app varied from 19 to 27 minutes. The coding of the Thinking Aloud method transcripts resulted in a total of 160 statements coded, including 44 positive aspects, 57 suggestions for improvement, 50 usability problems, and 9 functionality problems (all statements were counted individually). The coded transcript can be found in [Multimedia Appendix 5](#).

Positive Aspects

A total of 44 positive statements could be identified. As shown in [Table 2](#), these were categorized into six subcategories (grouped by participants).

Table 2. Distribution of the subcategories for “positive aspects.”

Subcategory (positive aspects)	Participants, n
Learnability	7
Navigation	7
Feedback	6
Instructions	5
Design	4
Structure	3

Functionality Problems

Two functionality problems were identified that were seen as most important. The first occurred while synchronizing the data collected offline. Two participants were unable to transfer the data. One participant commented: “‘Stop sending modified records to server’ [warning message in REDCap]—okay, let’s see what’s wrong. This should not happen, right?” Because the detailed description of the problem was in English, the participants could not solve it themselves. The problem arose because no data synchronization had occurred before the data collection. Therefore, the project in the app must be synchronized before the offline survey is carried out to prevent an interruption. The second problem became evident in the third task. The participant was supposed to conduct a follow-up interview in the record of the baseline interview (task 1). However, the participant created a new record and collected the

Seven out of 12 participants were optimistic about the app’s learnability and navigation: “Once you’ve done it a couple of times, you’re good with it” or “At the beginning, I was a bit confused...but now I understand it better already.” One research partner described the first steps from logging in to collecting data as “so even for me so far, foolproof.” Another positive aspect was the app’s feedback function for the project’s own built-in plausibility checks and warning messages: “Yes, that is also very helpful in any case, that you get the error message right away.” Feedback from the system such as receiving green symbols after successful saving of the questionnaires gave the participants a sense of security: “Yes and now it’s green. So, I assume that everything is saved.” In addition, some participants described the app as “clearly designed” and “well structured.”

follow-up interview data for a new study participant. The REDCap app should not offer this option as it requires further action to link the data to the existing data set from the baseline interview.

Usability Problems

Special attention was paid to the category “usability problems,” mainly influencing the app’s usability. A total of 50 statements were categorized into 6 subcategories. [Table 3](#) shows the usability problems, sorted by the severity score, which resulted from multiplying the severity rating by the number of mentions.

The REDCap app offers the option of translating the app’s interface using a language file. Nevertheless, not all terms have yet been translated into German. Some error and warning messages from the system are still in English, such as the message that appears when a user leaves a questionnaire without first saving it ([Figure 2](#)).

Table 3. Categories of usability problems and the derived importance discovered in the test.

Usability problem	Severity rating	Number of mentions	Severity score
Language	3	6	18
Feedback	3	4	12
Perceived offer character	2	6	12
Inconsistent interaction design	2	4	8
Navigation	1	7	7
Knowledge error	1	5	5

Figure 2. Screenshot of the REDCap app with a mix of German and English.



One participant was critical of the increased time required to interpret the English language: “I must ensure that I manage my time well and therefore I can’t constantly think about what that means.” The mixing of German and English bothered two users and resulted in some confusion: “It’s still a bit confusing. A few times it’s in German, then again in English. You have to switch quickly in your mind.” It is noticeable that the language barrier was complained about by participants from the two youngest age groups, who also achieved a lower SUS score.

The lack of information given to users about what is happening in the app (“feedback”) led to confusion during the survey and delayed task processing: “So here I’m not sure how to proceed.” This also led to uncertainty such as with regard to whether or not the synchronization task was successful: “For me, it’s unfortunately not apparent whether the data has been uploaded or not.” It is noteworthy that this usability problem only affected the participants who had no experience using tablets.

Participants were sometimes unaware of a function behind an interactive element such as a button (“perceived offer character”). For example, after logging into the app, the user must actively select a project to collect data, even if the user had been authorized only for one project. One participant described this situation as follows: “Next I go to ‘My Projects’ [4 second pause] Okay, the button’s missing, or I’m really having a blackout.” Especially in the older age group (45-59 years), buttons not being recognized as functions was a recurring problem.

In some situations, the participants expected a different function based on the design (“inconsistent interaction design”). The field type “date” was especially challenging for four participants. As shown in [Figure 3](#), the “date” field offers both a calendar icon and a textbox. However, the date could only be selected by clicking on the calendar icon. Four participants needed several attempts to select the correct date because the date field behaved differently from previous input fields.

Figure 3. Screenshot (in English and German) of a question with the field type “date”.



Although many participants appreciated the navigation in the app, they felt some navigation procedures to be too complicated (“navigation”). In particular, the pathway to logging out of the app was felt to be too cumbersome to find: “So how do I get out of here now?” Some usability problems were also caused by knowledge gaps (“knowledge error”): “How do I see what survey date I’ve chosen?”

Suggestions for Improvement

Some participants immediately provided suggestions for improvement after pointing out a problem. A total of 57 suggestions for improvement were identified, which were categorized into 5 subcategories (grouped by participants). Eight out of 12 participants suggested more explicit feedback, especially when synchronizing data during the second task: “Just a confirmation, for example, a pop-up message ‘Date successfully transferred,’ that I know I can go back.” Seven participants requested a “notes” field to collect additional data: “I would find it helpful to have the option to make notes quickly.” REDCap offers the function “field notes,” but it was not intuitive for the participants to find. Furthermore, 4 of the 8 participants would have found it more intuitive if they could have gone directly to the next question using the enter key: “For me, it would be helpful that it then jumps to the next question.” In addition, 5 out of 12 participants suggested an easier way to log out: “Simple ‘logout’ would have been clearer.” Four

participants proposed a more flexible option for language selection, either to switch entirely to German or to have more language options for colleagues whose native language is neither German nor English: “I also have colleagues who speak German but come from another country. So maybe another translation is necessary.” Moreover, 3 out of the 12 participants suggested more color tones and graphic highlighting in the design. Another participant would have preferred an indication of the progress of data collection: “for example, ‘You have completed 18 of 20 survey forms’.”

Identified Implications and Recommendations

Based on the usability problems found and suggestions for improvement, we identified implications and recommendations (Table 4). While analyzing the video and screen recordings, we observed that many of the usability problems led to uncertainty among participants, also resulting in delays in the survey process. In addition, this may cause data to be partially collected, or in the worst case, an interview could be cut short.

There are short-term solutions that the project team can provide, such as targeted user training for the identified issues or providing a test environment for users to familiarize themselves with the system. There is also long-term optimization potential that should be addressed by REDCap’s developers, such as including a user expertise-based help option within the app.

Table 4. Identified problems of the usability test, and short- or long-term solutions.

Problem	Short-term solution (by the project)	Long-term optimization (by the developers)
Usability problems		
Language	User training; provide a document with translations and explanations; check and adjust existing REDCap language file	Simplified language selection; complete translation, including system messages and synchronization report
Feedback	User training; provide a test system	User expertise-based help
Perceived offer character	User training; provide a test system	Optimization of user interface
Inconsistent interaction design	User training; provide a test system	Enable input by typing or hide input field; provide keyboard type based on the field type (eg, numeric keyboard)
Navigation	User training; create a short paper-based how-to guide	Simplification of navigation (eg, log out)
Knowledge error	User training; provide filling out instructions directly in the questionnaire	User expertise-based help
Functionality problems		
Data synchronization	User training; include a note at the end of the questionnaire to regularly synchronize the records	Sending notifications on the device in case of nonsynchronized records
Follow-up interview	User training; create a short paper-based how-to guide for follow-up interviews	Optimized visit overview

SUS Questionnaire

The overall SUS score of the REDCap app was 74. According to the Bangor classification, this represents “good usability” [44]. The SUS scores for each participant are shown in Table 5. The rows represent the recoded data of individual SUS questions and the columns refer to the individual participants. The highest value was 90 and the lowest value was 55. Participants from the oldest and the second-youngest age groups rated the app with “excellent” usability (SUS=90). Participants

with experience using a tablet achieved an average SUS score of 75.8, compared to 72.9 for participants who did not have tablet experience.

The majority of participants indicated that they would use the REDCap app frequently (8 of 12 participants). Furthermore, 8 participants found the app easy to use. Ten out of 12 participants could imagine that most people can quickly learn to use REDCap. Concerning the need for technical assistance in using the app, the results indicated good usability, as 10 of the 12

respondents disagreed or strongly disagreed that help was needed. Nevertheless, half of the participants felt the system was unnecessarily complex and 5 participants disagreed with

the statement about feeling confident using the app. None of the negatively worded questions received a “strongly agree” rating.

Table 5. Detailed System Usability Scale (SUS) scores for each participant (N=12).

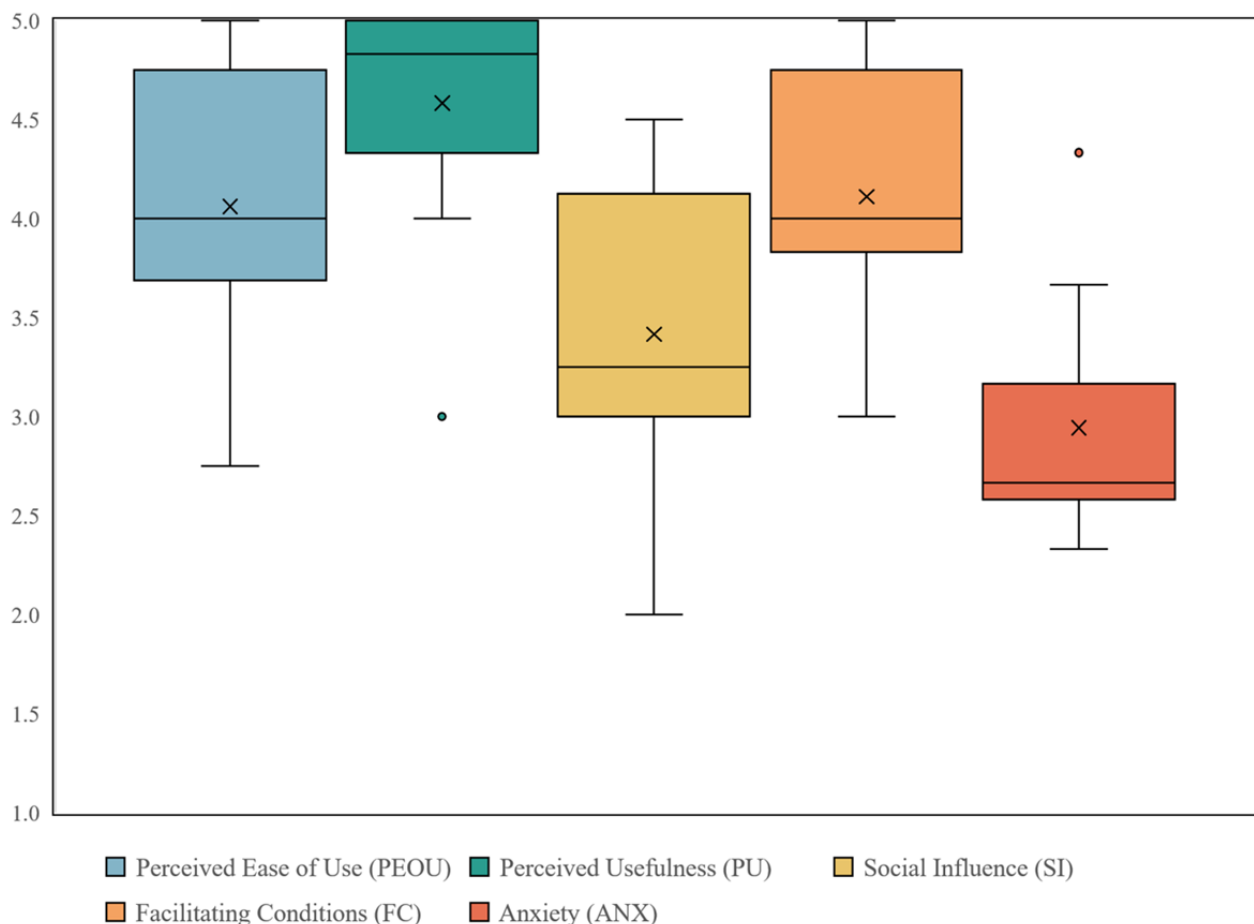
SUS item	Participant number (age group, years)												Mean (SD)	Total	SUS score
	P1 (25-34)	P2 (18-24)	P3 (45-59)	P4 (45-59)	P5 (45-59)	P6 (25-34)	P7 (25-34)	P8 (35-44)	P9 (25-34)	P10 (45-59)	P11 (>60)	P12 (45-59)			
I think that I would like to use this system frequently	2	2	3	2	3	3	4	2	3	3	4	4	2.91 (0.79)	35.0	87.5
I found the system unnecessarily complex	2	3	2	1	2	3	4	4	4	1	4	3	2.75 (1.13)	33.0	82.5
I thought the system was easy to use	3	2	4	2	1	4	3	2	3	3	3	3	2.75 (0.86)	33.0	82.5
I think that I would need the support of a technical person to be able to use this system	3	3	4	3	1	4	4	3	3	1	4	3	3.00 (1.04)	36.0	90.0
I found the various functions in this system were well integrated	3	2	4	2	4	3	3	2	4	3	3	4	3.08 (0.79)	37.0	92.5
I thought there was too much inconsistency in this system	2	4	2	4	2	3	3	3	4	1	3	3	2.83 (0.75)	34.0	85.0
I would imagine that most people would learn to use this system very quickly	3	3	4	4	3	4	3	2	2	4	3	4	3.25 (0.75)	39.0	97.5
I found the system very cumbersome to use	2	3	4	3	2	4	3	3	4	4	4	4	3.33 (0.77)	40.0	100
I felt very confident using the system	2	1	3	1	2	4	3	2	3	4	4	3	2.66 (1.07)	32.0	80.0
I needed to learn a lot of things before I could get going with this system	2	3	2	4	2	4	3	3	3	4	4	4	3.16 (0.83)	38.0	95.0
Total	24	26	32	26	22	36	33	26	33	28	36	35	N/A ^a	N/A	N/A
SUS score	60	65	80	65	55	90	82.5	65	82.5	70	90	87.5	N/A	N/A	N/A

^aN/A: not applicable.

Technology Acceptance Questionnaire

An extended TAM was used to evaluate how the participants accept mobile devices for data collection. From a maximum 75-point scale (15 questions multiplied by the highest response value of 5), the participants scored an average of 60.5 points (SD 7.3). The number of items collected in the questionnaire was reduced to the following underlying factors that determine the items' average scale value: perceived ease of use, perceived usefulness, social influence, facilitating conditions, and anxiety. The boxplot in Figure 4 shows the distribution of the technology acceptance categories. The perceived usefulness of mobile

devices was rated as very high by participants (mean 4.58), indicating that they consider mobile devices to be helpful in their work, as these enable faster processing of tasks and make work easier. The category anxiety met with the lowest approval (mean 2.94), which indicates that the participants have little concern about using mobile devices and are not afraid to use them for their work. The largest heterogeneity in responses (SD 0.82) was found in the category social influence, which implies that the participants tend to receive different recommendations on the use of mobile devices from their work colleagues and that some supervisors support the use of mobile devices while others do not.

Figure 4. Distribution of the Technology Acceptance Model categories.

Discussion

Principal Results

The success of a registry depends on the quality of the data collected. Most registry studies use pilot testing to evaluate the correct implementation of the questionnaires used in the EDC system. A usability evaluation of whether users can cope with the EDC system in the intended environment is often ignored. This is even more important because usability problems can affect whether an app is ultimately adopted or abandoned [11,72].

Most of the digiDEM partners had not yet had any experience in registry research. They are therefore considered to be “lay users,” as they are neither familiar with registry research studies nor using an app on a mobile device for data collection. Conventional EDC systems are intended to be used by professional registry research staff at a clinic site [22]. In addition, there are numerous differences between web-based and app-based data collection that can affect usability, such as smaller displays with adjusted interfaces and different input devices on mobile devices or different workflows due to data synchronization. In longitudinal studies such as digiDEM Bayern, it is essential to adhere as closely as possible to the specified survey dates to guarantee the study’s timeliness and validity [21]. To avoid waiting until a patient returns to a particular facility or environment, offline on-site data collection

via a mobile app provides a valuable tool for more effective decentralized registry studies.

Our usability evaluation helped us to identify issues that could affect the usability of offline data collection with the REDCap app (Table 4). With the help of the participants’ thought processes during the Thinking Aloud test, it was possible to gain a more nuanced understanding of participants’ behaviors, intentions, and expectations. Therefore, we were able to uncover usability issues even after participants had successfully completed a task. Participants also made suggestions for improvement to increase efficiency and data quality. The three most significant usability issues were the language barrier, lack of feedback, and unperceived nature of the procedures the app offered. Although these are not critical problems, they do limit usability.

The language by which a system communicates with a user can have a major influence on usability [22,73]. For the efficient use of an EDC system, inexperienced users should especially be able to use their native language. The language could affect the acceptance of a system, as shown in other studies [74,75]. Participants’ statements in our study confirmed this. In contrast to English system messages, messages in German (eg, those arising from self-designed plausibility checks) were perceived as positive by the participants. Bearing this in mind, a complete translation of the app should lead to an increased SUS score (ie, a higher rating of the usability of the REDCap app).

In their study about the impact of usability on software design, Juristo et al [73] found further usability issues that were also addressed by our participants, such as “feedback.” Insufficient feedback from the REDCap app, identified as a usability issue by some participants, made participants feel insecure.

All field types used in the questionnaire should be tested, especially when using nontext-based items such as date selection fields or visual scales [76]. To ensure that users are able to complete the study questionnaire, these nontext-based items must be as usable on the tablet as regular text-based items.

The data synchronization process was identified as a major functionality usability problem. Uploading the data collected offline was also noted as problematic in a study of implementation strategies for the REDCap app by McIntosh et al [77]. The developers of REDCap acknowledge [30]: “Sending data is, at present, a complicated process.” Since the app does not automatically upload the collected records as soon as an internet connection is available, the user must actively start the process. This can lead to unsynced data and data loss in the worst case. Therefore, the process and its importance must be explicitly explained and trained.

Among the positive aspects of the qualitative usability evaluation, learnability was particularly highlighted. The participants who were not experienced in handling mobile devices became more familiar with the app and the tablet from task to task. The results of the quantitative SUS questionnaire confirmed the positive statements. For example, learnability was also one of the highest-rated items in the SUS. Given the widespread use of the SUS, a comparison with existing SUS study results is possible [78]. With a SUS score of 74, the REDCap app’s usability is above the average SUS score of 68, which was calculated from the results of around 500 studies [44].

Good usability cannot always predict the likelihood of future use, as other factors may also play a role. For example, anxiety can lead to a system being perceived as not easy to use, even if it has been designed to be user-friendly [54]. It is essential to ensure that future users accept an app on a mobile device for data collection before introducing such technology. In addition to the costs involved, less user acceptance of the technology can lead to less frequent use [79] and lower performance [80]. For these reasons, eliciting feedback from data collectors is essential to avoid jeopardizing success in a registry project. Considering that many studies have problems reaching their target number of participants [81,82], a lack of technology acceptance or usability issues should not additionally lead to lower numbers of participants. The results of our TAM questionnaire support the findings of the Thinking Aloud test and the SUS questionnaire that the REDCap app and tablets will be accepted and used by the user group examined here. The two categories “perceived usefulness” and “perceived ease of use” were rated highly. “Anxiety” can also be ruled out as an obstacle to using the technology. Only in the area of “social influence” would broader support be desirable. Here, the lack of digitization at some institutions is a likely variable [83,84].

By evaluating the usability and acceptance of app-based offline data collection at an early stage in our project, we were able to

identify usability problems that need to be considered when introducing such a data collection method. As Qiu and Yu [85] have shown, the results of the Thinking Aloud test can be used not only to evaluate usability but also to determine the training needs of novice system users. We were able to confirm this with our findings (see Table 4). Based on the identified issues, individual training measures focusing on certain issues could be generated. Furthermore, setting up a test environment should help reduce participants’ feelings of insecurity before an initial survey.

Comparison With Prior Work

The advantages of a system with good usability, such as enhanced efficiency and user acceptance, less training effort, or higher data quality, are indisputable [11,86]. However, usability evaluations are rarely used for EDC systems in clinical and registry research. Welker [87] identified numerous barriers and solutions during the implementation of EDC systems without explicitly addressing the usability issue. McIntosh et al [77] described an implementation strategy of mobile technologies for data collection using the REDCap app in survey research projects. The primary focus of their study was the technical feasibility of the offline data collection process rather than its usability. There are also guidance documents in place [88] or under development [89] that address the requirements and issues of EDC in clinical trials. However, usability is not, or only briefly, mentioned in these guidelines.

Only a few studies such as those by Ndlovu et al [90] and Dunn et al [91] have compared the usability of the EDC system REDCap to other data collection methods such as paper-based or Microsoft Excel/Access. Due to the limitation of using only a quantitative questionnaire, they could not identify specific usability issues in their studies. To our knowledge, there are no other studies that feature a usability evaluation of the REDCap mobile data collection using a qualitative-driven mixed methods approach. This also allowed us to identify specific usability problems.

Walden et al [22] recently took a different approach, suggesting that it is useful to perform a usability study while EDC tools are still being developed to improve the usability of such systems from the start. Nevertheless, the results of our work suggest that registry studies with already developed EDC systems can benefit from usability evaluations with respect to the individual registry requirements and the particular health care setting.

Limitations

Three limitations of the study should be acknowledged. First, the sample size (N=12) could be seen as too small for meaningful assessments to be extrapolated. Even though larger samples are usually recommended for quantitative studies, Tullis and Stetson [46] and Nielsen and Landauer [92] have shown that the SUS questionnaire and usability tests can deliver meaningful results even with 12 participants. Furthermore, Nielsen [93] and Faulkner [94] confirmed that a small sample of test subjects is sufficient to apply the Thinking Aloud method as part of a usability test. Compared to other studies, similarly large sample sizes could provide valid results [95]. The sample

size also depends on how many different platforms are evaluated [61]. Therefore, it can be assumed that 12 participants was a suitable sample size for this study.

Second, participation in the study was voluntary. In this case, only participants who had experience with mobile devices or technically interested participants can be presumed to have participated in the study. As shown by the participants' characteristics (Table 1), only half the sample had experience using tablets. This ratio corresponds to the German average [96]. One participant had no mobile device experience at all. However, it cannot be ruled out that usability could be rated more highly in a sample of participants who are familiar with mobile devices.

Third, the usability testing was laboratory-based. We replicated a real scenario [97] for all participants by simulating the context and the interviewing of people with dementia. Simulating an interview partner is essential for systems that use branching logic, where the order of the questions is determined by the interview partner's answer [98]. Nevertheless, it cannot be excluded that the app's usability during actual interviews of people with dementia might differ from our results [21].

Conclusions

Offline registry data collection can be made more efficient through EDC systems, but attention must be paid to the usability

of these systems. Despite the widespread use of usability tests in the health care and app environment, usability evaluations in the field of electronic data collection in registry-based research have so far remained scarce. Our study shows that it is profitable to conduct a usability evaluation of the EDC system considering future users and the project environment. Using a mixed method approach, we identified positive and negative aspects regarding the usability of an EDC app for offline data collection. By addressing these aspects, the registry project digiDEM Bayern can avoid pitfalls and realize the benefits of EDC systems, even in areas where using web-based EDC systems is not viable due to unreliable internet connectivity. The out-of-the-box use of the REDCap app resulted in a good usability rating, which can be further improved by addressing the identified issues by means of user training of digiDEM partners and improvements on the part of REDCap's developers. The technology acceptance in the sample showed that heterogeneous groups of different ages with varying experiences in handling mobile devices are open to the use of app-based EDC systems. Based on these results, it can be assumed that the offline use of an app-based EDC system on mobile devices is a viable solution for collecting data in a registry-based research project.

Acknowledgments

This study was performed in (partial) fulfillment of the requirements for obtaining the degree "Dr. rer. biol. hum." from the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) (MR). The project digiDEM Bayern is funded by the Bavarian State Ministry of Health and Care as part of the funding initiative "BAYERN DIGITAL II" (funding code G42d-G8300-2017/1606-83).

Authors' Contributions

MR and MH planned, designed, and performed the usability evaluation. MR assisted during the design and evaluation phases. MR advised in the preparation phase of the study and assisted in collecting the evaluation data. MR coordinated input from the coauthors. MR participated in writing the paper. EG, PR, HP, and MR revised the first draft and provided valuable input and comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Test survey in the REDCap mobile app.

[\[PDF File \(Adobe PDF File\), 2592 KB - formative_v5i11e31649_app1.pdf\]](#)

Multimedia Appendix 2

Test manual with tasks and answers for the test survey.

[\[PDF File \(Adobe PDF File\), 690 KB - formative_v5i11e31649_app2.pdf\]](#)

Multimedia Appendix 3

Introductory training to the REDCap mobile app.

[\[PDF File \(Adobe PDF File\), 1413 KB - formative_v5i11e31649_app3.pdf\]](#)

Multimedia Appendix 4

Online questionnaire with sociodemographic data, System Usability Scale, and technology acceptance model.

[\[PDF File \(Adobe PDF File\), 328 KB - formative_v5i11e31649_app4.pdf\]](#)

Multimedia Appendix 5

Thinking Aloud results (encoded segments).

[\[PDF File \(Adobe PDF File\), 613 KB - formative_v5i11e31649_app5.pdf\]](#)

References

1. Richesson R, Vehik K. Patient registries: utility, validity and inference. *Adv Exp Med Biol* 2010;686:87-104. [doi: [10.1007/978-90-481-9485-8_6](https://doi.org/10.1007/978-90-481-9485-8_6)] [Medline: [20824441](https://pubmed.ncbi.nlm.nih.gov/20824441/)]
2. Gliklich RE, Dreyer NA, Leavy MB, Christian JB. 21st Century Patient Registries. EBook addendum to *Registries for Evaluating Patient Outcomes: A User's Guide*, 3rd Edition. In: AHRQ Publication No. 17(18)-EHC013-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2018.
3. Zozus M, Sanns B, Eisenstein E. Beyond EDC. *J Soc Clin Data Manag* 2021;1(1):1-22. [doi: [10.47912/jscdm.33](https://doi.org/10.47912/jscdm.33)]
4. Lallas CD, Preminger GM, Pearle MS, Leveillee RJ, Lingeman JE, Schwoppe JP, et al. Internet based multi-institutional clinical research: a convenient and secure option. *J Urol* 2004 May;171(5):1880-1885. [doi: [10.1097/01.ju.0000120221.39184.3c](https://doi.org/10.1097/01.ju.0000120221.39184.3c)] [Medline: [15076298](https://pubmed.ncbi.nlm.nih.gov/15076298/)]
5. Paul J, Seib R, Prescott T. The internet and clinical trials: background, online resources, examples and issues. *J Med Internet Res* 2005 Mar 16;7(1):e5 [FREE Full text] [doi: [10.2196/jmir.7.1.e5](https://doi.org/10.2196/jmir.7.1.e5)] [Medline: [15829477](https://pubmed.ncbi.nlm.nih.gov/15829477/)]
6. Neuer A, Warnock N, Slezinger E. The upfront cost hurdle of EDC. *Appl Clin Trial* 2010;19(4):58-54.
7. Staziaki PV, Kim P, Vadvala HV, Ghoshhajra BB. Medical registry data collection efficiency: a crossover study comparing web-based electronic data capture and a standard spreadsheet. *J Med Internet Res* 2016 Jun 08;18(6):e141 [FREE Full text] [doi: [10.2196/jmir.5576](https://doi.org/10.2196/jmir.5576)] [Medline: [27277523](https://pubmed.ncbi.nlm.nih.gov/27277523/)]
8. Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. *PLoS One* 2011;6(9):e25348 [FREE Full text] [doi: [10.1371/journal.pone.0025348](https://doi.org/10.1371/journal.pone.0025348)] [Medline: [21966505](https://pubmed.ncbi.nlm.nih.gov/21966505/)]
9. Malik I, Burnett S, Webster-Smith M, Morden J, Ereira S, Gillman A, et al. Benefits and challenges of electronic data capture (EDC) systems versus paper case report forms. *Trials* 2015 Nov 16;16(S2):37. [doi: [10.1186/1745-6215-16-s2-p37](https://doi.org/10.1186/1745-6215-16-s2-p37)]
10. Fleischmann R, Decker A, Kraft A, Mai K, Schmidt S. Mobile electronic versus paper case report forms in clinical trials: a randomized controlled trial. *BMC Med Res Methodol* 2017 Dec 01;17(1):153 [FREE Full text] [doi: [10.1186/s12874-017-0429-y](https://doi.org/10.1186/s12874-017-0429-y)] [Medline: [29191176](https://pubmed.ncbi.nlm.nih.gov/29191176/)]
11. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform* 2009 May;78(5):340-353. [doi: [10.1016/j.ijmedinf.2008.10.002](https://doi.org/10.1016/j.ijmedinf.2008.10.002)] [Medline: [19046928](https://pubmed.ncbi.nlm.nih.gov/19046928/)]
12. Hoerbst A, Schweitzer M. A systematic investigation on barriers and critical success factors for clinical information systems in integrated care settings. *Yearb Med Inform* 2015 Aug 13;10(1):79-89 [FREE Full text] [doi: [10.15265/IY-2015-018](https://doi.org/10.15265/IY-2015-018)] [Medline: [26293853](https://pubmed.ncbi.nlm.nih.gov/26293853/)]
13. Meadows BJ. Eliciting remote data entry system requirements for the collection of cancer clinical trial data. *Comput Inform Nurs* 2003;21(5):234-240. [doi: [10.1097/00024665-200309000-00010](https://doi.org/10.1097/00024665-200309000-00010)] [Medline: [14504599](https://pubmed.ncbi.nlm.nih.gov/14504599/)]
14. Szajna B. Empirical evaluation of the revised technology acceptance model. *Manag Sci* 1996 Jan;42(1):85-92. [doi: [10.1287/mnsc.42.1.85](https://doi.org/10.1287/mnsc.42.1.85)]
15. Johnson CM, Johnston D, Crowley PK, Culbertson H, Rippen HE, Damico DJ, et al. EHR Usability Toolkit: A Background Report on Usability and Electronic Health Records (Prepared by Westat under Contract No. HHS 290-2009-00023I). In: AHRQ Publication No. 11-0084-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2011.
16. Wronikowska MW, Malycha J, Morgan LJ, Westgate V, Petrinic T, Young JD, et al. Systematic review of applied usability metrics within usability evaluation methods for hospital electronic healthcare record systems: Metrics and Evaluation Methods for eHealth Systems. *J Eval Clin Pract* 2021 May 13:online ahead of print. [doi: [10.1111/jep.13582](https://doi.org/10.1111/jep.13582)] [Medline: [33982356](https://pubmed.ncbi.nlm.nih.gov/33982356/)]
17. Broekhuis M, van Velsen L, Hermens H. Assessing usability of eHealth technology: A comparison of usability benchmarking instruments. *Int J Med Inform* 2019 Aug;128:24-31. [doi: [10.1016/j.ijmedinf.2019.05.001](https://doi.org/10.1016/j.ijmedinf.2019.05.001)] [Medline: [31160008](https://pubmed.ncbi.nlm.nih.gov/31160008/)]
18. Al-Saadi TA, Aljarrah TM, Alhashemi AM, Hussain A. A systematic review of usability challenges and testing in mobile health. *Int J Account Financ Report* 2015 Jul 12;5(2):1. [doi: [10.5296/ijaf.v5i2.8004](https://doi.org/10.5296/ijaf.v5i2.8004)]
19. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: systematic review. *J Med Internet Res* 2020 Jan 06;22(1):e15337 [FREE Full text] [doi: [10.2196/15337](https://doi.org/10.2196/15337)] [Medline: [31904579](https://pubmed.ncbi.nlm.nih.gov/31904579/)]
20. Patel B, Thind A. Usability of mobile health apps for postoperative care: systematic review. *JMIR Perioper Med* 2020 Jul 20;3(2):e19099 [FREE Full text] [doi: [10.2196/19099](https://doi.org/10.2196/19099)] [Medline: [33393925](https://pubmed.ncbi.nlm.nih.gov/33393925/)]
21. Schmier JK, Kane DW, Halpern MT. Practical applications of usability theory to electronic data collection for clinical trials. *Contemp Clin Trials* 2005 Jun;26(3):376-385. [doi: [10.1016/j.cct.2005.01.012](https://doi.org/10.1016/j.cct.2005.01.012)] [Medline: [15911471](https://pubmed.ncbi.nlm.nih.gov/15911471/)]
22. Walden A, Garvin L, Smerek M, Johnson C. User-centered design principles in the development of clinical research tools. *Clin Trials* 2020 Dec;17(6):703-711. [doi: [10.1177/1740774520946314](https://doi.org/10.1177/1740774520946314)] [Medline: [32815381](https://pubmed.ncbi.nlm.nih.gov/32815381/)]
23. Dietzel N, Kürten L, Karrer L, Reichold M, Köhler L, Nagel A, et al. Digital dementia registry Bavaria-digiDEM Bayern: study protocol for a multicentre, prospective, longitudinal register study. *BMJ Open* 2021 Feb 08;11(2):e043473 [FREE Full text] [doi: [10.1136/bmjopen-2020-043473](https://doi.org/10.1136/bmjopen-2020-043473)] [Medline: [33558357](https://pubmed.ncbi.nlm.nih.gov/33558357/)]

24. Reichold M, Dietzel N, Chmelirsch C, Kolominsky-Rabas PL, Graessel E, Prokosch H. Designing and implementing an IT architecture for a digital multicenter dementia registry: digiDEM Bayern. *Appl Clin Inform* 2021 May;12(3):551-563 [FREE Full text] [doi: [10.1055/s-0041-1731286](https://doi.org/10.1055/s-0041-1731286)] [Medline: [34134149](https://pubmed.ncbi.nlm.nih.gov/34134149/)]
25. 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/)]
26. Tolea MI, Morris JC, Galvin JE. Trajectory of mobility decline by type of dementia. *Alzheimer Dis Assoc Disord* 2016;30(1):60-66 [FREE Full text] [doi: [10.1097/WAD.0000000000000091](https://doi.org/10.1097/WAD.0000000000000091)] [Medline: [25886717](https://pubmed.ncbi.nlm.nih.gov/25886717/)]
27. Sylliaas H, Selbaek G, Bergland A. Do behavioral disturbances predict falls among nursing home residents? *Aging Clin Exp Res* 2012 Jun;24(3):251-256. [doi: [10.1007/BF03325253](https://doi.org/10.1007/BF03325253)] [Medline: [23114551](https://pubmed.ncbi.nlm.nih.gov/23114551/)]
28. Huxhold O, Otte K. Internetzugang und Internetnutzung in der zweiten Lebenshälfte. In: *dza aktuell Deutscher Alterssurvey*. Berlin: Deutsches Zentrum für Altersfragen; 2019.
29. Initiative D21. D21-Digital-Index 2018 / 2019 - Jährliches Lagebild zur Digitalen Gesellschaft. Berlin: Initiative D21 e. V; 2019.
30. REDCap Mobile App Guide. Vanderbilt University. 2016. URL: <https://projectredcap.org/wp-content/uploads/2016/08/REDCap-Mobile-App-Guide.docx> [accessed 2021-08-22]
31. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
32. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. *Kolner Z Soz Sozpsychol* 2017;69(Suppl 2):107-131 [FREE Full text] [doi: [10.1007/s11577-017-0454-1](https://doi.org/10.1007/s11577-017-0454-1)] [Medline: [28989188](https://pubmed.ncbi.nlm.nih.gov/28989188/)]
33. Someren MV, Barnard Y, Sandberg J. The think aloud method: a practical guide to modelling cognitive processes. London: Academic Press; 1994.
34. Leech NL, Onwuegbuzie AJ. A typology of mixed methods research designs. *Qual Quant* 2007 Mar 27;43(2):265-275. [doi: [10.1007/s11135-007-9105-3](https://doi.org/10.1007/s11135-007-9105-3)]
35. Curry L, Nunez-Smith M. Mixed methods in health sciences research: a practical primer. Los Angeles: SAGE Publications; 2015.
36. Carayon P, Kianfar S, Li Y, Xie A, Alyousef B, Wooldridge A. A systematic review of mixed methods research on human factors and ergonomics in health care. *Appl Ergon* 2015 Nov;51:291-321 [FREE Full text] [doi: [10.1016/j.apergo.2015.06.001](https://doi.org/10.1016/j.apergo.2015.06.001)] [Medline: [26154228](https://pubmed.ncbi.nlm.nih.gov/26154228/)]
37. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. *Int J Med Inform* 2019 Jun;126:95-104. [doi: [10.1016/j.ijmedinf.2019.03.018](https://doi.org/10.1016/j.ijmedinf.2019.03.018)] [Medline: [31029270](https://pubmed.ncbi.nlm.nih.gov/31029270/)]
38. Sarodnick F, Brau H. Methoden der usability evaluation: wissenschaftliche grundlagen und praktische anwendung. 2., überarb. und aktualisierte Aufl. Bern: Huber; 2011.
39. Nielsen J, Clemmensen T, Yssing C. Getting access to what goes on in people's heads?: reflections on the think-aloud technique. USA: Association for Computing Machinery; 2002 Presented at: NordiCHI '02: Second Nordic Conference on Human-Computer Interaction; October 19-23, 2002; Aarhus, Denmark p. 101-110. [doi: [10.1145/572020.572033](https://doi.org/10.1145/572020.572033)]
40. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. *IEEE Trans Profess Commun* 2000;43(3):261-278. [doi: [10.1109/47.867942](https://doi.org/10.1109/47.867942)]
41. Nisbett RE, Wilson TD. Telling more than we can know: Verbal reports on mental processes. *Psychol Rev* 1977 May;84(3):231-259. [doi: [10.1037/0033-295x.84.3.231](https://doi.org/10.1037/0033-295x.84.3.231)]
42. Leiner D. SoSci Survey (Version 3). URL: <https://www.soscisurvey.de> [accessed 2021-08-22]
43. Brooke J. SUS: A "Quick and Dirty" Usability Scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability evaluation in industry. London: CRC Press; 1996:207-212.
44. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. *Int J Hum-Comput Interact* 2008 Jul 30;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
45. Lewis JR, Sauro J. The factor structure of the System Usability Scale. Berlin Heidelberg; 2009 Presented at: HCD 2009 Human Centered Design; July 19-24, 2009; San Diego, CA p. 94-103. [doi: [10.1007/978-3-642-02806-9_12](https://doi.org/10.1007/978-3-642-02806-9_12)]
46. Tullis T, Stetson J. A comparison of questionnaires for assessing website usability. 2004 Presented at: Usability Professionals Association (UPA); June 7-11, 2004; Minneapolis. [doi: [10.1007/978-1-84628-795-4_3](https://doi.org/10.1007/978-1-84628-795-4_3)]
47. Gao M, Kortum P, Oswald FL. Multi-Language Toolkit for the System Usability Scale. *International Journal of Human-Computer Interaction* 2020;36(20):1883-1901. [doi: [10.1080/10447318.2020.1801173](https://doi.org/10.1080/10447318.2020.1801173)]
48. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quart* 1989 Sep;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
49. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. *Appl Clin Inform* 2018 Jul;9(3):604-634 [FREE Full text] [doi: [10.1055/s-0038-1668091](https://doi.org/10.1055/s-0038-1668091)] [Medline: [30112741](https://pubmed.ncbi.nlm.nih.gov/30112741/)]
50. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010 Feb;43(1):159-172 [FREE Full text] [doi: [10.1016/j.jbi.2009.07.002](https://doi.org/10.1016/j.jbi.2009.07.002)] [Medline: [19615467](https://pubmed.ncbi.nlm.nih.gov/19615467/)]

51. Ammenwerth E. Technology acceptance models in health informatics: TAM and UTAUT. *Stud Health Technol Inform* 2019 Jul 30;263:64-71. [doi: [10.3233/SHTI190111](https://doi.org/10.3233/SHTI190111)] [Medline: [31411153](https://pubmed.ncbi.nlm.nih.gov/31411153/)]
52. Saade R, Kira D. The emotional state of technology acceptance. 2006 Presented at: InSITE 2006: Informing Science + IT Education Conference; June 25-28, 2006; Manchester. [doi: [10.28945/2945](https://doi.org/10.28945/2945)]
53. Dönmez-Turan A, Kir M. User anxiety as an external variable of technology acceptance model: A meta-analytic study. *Proc Comput Sci* 2019;158:715-724. [doi: [10.1016/j.procs.2019.09.107](https://doi.org/10.1016/j.procs.2019.09.107)]
54. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quart* 2012;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
55. 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)]
56. Lee Y, Lee J, Lee Z. Social influence on technology acceptance behavior. *SIGMIS Database* 2006 Sep 19;37(2-3):60-75. [doi: [10.1145/1161345.1161355](https://doi.org/10.1145/1161345.1161355)]
57. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Quart* 2003;27(3):425. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
58. Spencer L, Ritchie J, Lewis J, Dillon L. Quality in qualitative evaluation: a framework for assessing research evidence - a quality framework. London: Government Chief Social Researcher's Office; 2003.
59. Sandelowski M. Sample size in qualitative research. *Res Nurs Health* 1995 Apr;18(2):179-183. [doi: [10.1002/nur.4770180211](https://doi.org/10.1002/nur.4770180211)] [Medline: [7899572](https://pubmed.ncbi.nlm.nih.gov/7899572/)]
60. Sonderegger A, Schmutz S, Sauer J. The influence of age in usability testing. *Appl Ergon* 2016 Jan;52:291-300. [doi: [10.1016/j.apergo.2015.06.012](https://doi.org/10.1016/j.apergo.2015.06.012)] [Medline: [26360221](https://pubmed.ncbi.nlm.nih.gov/26360221/)]
61. Geisen E, Bergstrom J. Usability testing for survey research. Cambridge, MA: Morgan Kaufmann Publishers; Feb 15, 2017.
62. Nielsen J. Usability inspection methods. 1994 Presented at: Conference Companion on Human Factors in Computing Systems - CHI '94; April 24-28, 1994; Boston.
63. Nielsen J. Usability engineering. London: Academic Press; 1993.
64. Mayring P. Qualitative inhaltsanalyse. In: Mey G, Mruck K, editors. *Handbuch qualitative forschung in der psychologie*. Wiesbaden: VS Verlag für Sozialwissenschaften; 2010:601-613.
65. Elo S, Kääriäinen M, Kanste O, Pölkki T, Utriainen K, Kyngäs H. Qualitative content analysis: a focus on trustworthiness. *SAGE Open* 2014 Feb 11;4(1):215824401452263. [doi: [10.1177/2158244014522633](https://doi.org/10.1177/2158244014522633)]
66. Kuckartz U. *Qualitative inhaltsanalyse: methoden, praxis, computerunterstützung*. Weinheim Basel: Beltz-Juventa; 2018.
67. Zapf D, Brodbeck F, Prümper J. Handlungsorientierte fehlertaxonomie in der mensch-computer interaktion. *Zeitschrift für Arbeits- und Organisationspsychologie* 1989;33(4):178-187.
68. Prümper J. Fehlerbeurteilungen in der mensch-computer interaktion: reliabilitätsanalysen und training einer handlungstheoretischen fehlertaxonomie. Münster: Waxmann; 1994.
69. Rödiger KH. Anwendungsbereiche lernen voneinander ... und woraus lernen wir? In: *Software-Ergonomie '95*. Stuttgart: Vieweg+Teubner; 1995:43-60.
70. Noble H, Heale R. Triangulation in research, with examples. *Evid Based Nurs* 2019 Jul;22(3):67-68. [doi: [10.1136/ebnurs-2019-103145](https://doi.org/10.1136/ebnurs-2019-103145)] [Medline: [31201209](https://pubmed.ncbi.nlm.nih.gov/31201209/)]
71. Wicks D. The coding manual for qualitative researchers (3rd edition). *Res Organiz Manag* 2017 Jun 12;12(2):169-170. [doi: [10.1108/qrom-08-2016-1408](https://doi.org/10.1108/qrom-08-2016-1408)]
72. 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/)]
73. Juristo N, Moreno AM, Sanchez-Segura M. Analysing the impact of usability on software design. *J Syst Soft* 2007 Sep;80(9):1506-1516. [doi: [10.1016/j.jss.2007.01.006](https://doi.org/10.1016/j.jss.2007.01.006)]
74. Ross J, Gao J. Overcoming the language barrier in mobile user interface design: A case study on a mobile health app. 2015 Presented at: Australasian Conference on Information Systems; 2015; Adelaide.
75. Nantel J, Glaser E. The impact of language and culture on perceived website usability. *J Engineer Technol Manag* 2008 Mar;25(1-2):112-122. [doi: [10.1016/j.jengtecman.2008.01.005](https://doi.org/10.1016/j.jengtecman.2008.01.005)]
76. Jamison RN, Gracely RH, Raymond SA, Levine JG, Marino B, Herrmann TJ, et al. Comparative study of electronic vs. paper VAS ratings: a randomized, crossover trial using healthy volunteers. *Pain* 2002 Sep;99(1-2):341-347. [doi: [10.1016/s0304-3959\(02\)00178-1](https://doi.org/10.1016/s0304-3959(02)00178-1)] [Medline: [12237213](https://pubmed.ncbi.nlm.nih.gov/12237213/)]
77. McIntosh S, Pérez-Ramos J, Demment MM, Vélez Vega C, Avedaño E, Ossip DJ, et al. Development and implementation of culturally tailored offline mobile health surveys. *JMIR Public Health Surveill* 2016 Jun 02;2(1):e28 [FREE Full text] [doi: [10.2196/publichealth.5408](https://doi.org/10.2196/publichealth.5408)] [Medline: [27256208](https://pubmed.ncbi.nlm.nih.gov/27256208/)]
78. Lewis JR. The System Usability Scale: past, present, and future. *Int J Hum-Comput Interact* 2018 Mar 30;34(7):577-590. [doi: [10.1080/10447318.2018.1455307](https://doi.org/10.1080/10447318.2018.1455307)]
79. Turner M, Kitchenham B, Brereton P, Charters S, Budgen D. Does the technology acceptance model predict actual use? A systematic literature review. *Inf Soft Technol* 2010 May;52(5):463-479. [doi: [10.1016/j.infsof.2009.11.005](https://doi.org/10.1016/j.infsof.2009.11.005)]
80. Devaraj S, Kohli R. Performance impacts of information technology: is actual usage the missing link? *Manag Sci* 2003 Mar;49(3):273-289. [doi: [10.1287/mnsc.49.3.273.12736](https://doi.org/10.1287/mnsc.49.3.273.12736)]

81. Sully BGO, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. *Trials* 2013 Jun 09;14:166 [FREE Full text] [doi: [10.1186/1745-6215-14-166](https://doi.org/10.1186/1745-6215-14-166)] [Medline: [23758961](https://pubmed.ncbi.nlm.nih.gov/23758961/)]
82. Walters SJ, Bonacho Dos Anjos Henriques-Cadby I, Bortolami O, Flight L, Hind D, Jacques RM, et al. Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme. *BMJ Open* 2017 Mar 20;7(3):e015276 [FREE Full text] [doi: [10.1136/bmjopen-2016-015276](https://doi.org/10.1136/bmjopen-2016-015276)] [Medline: [28320800](https://pubmed.ncbi.nlm.nih.gov/28320800/)]
83. Behm M, Klenk T. Digitalisierung im Gesundheitssektor. Wiesbaden: Springer Fachmedien; 2019.
84. Caumanns J. For discussion: The state of digitization of the German healthcare system. *Z Evid Fortbild Qual Gesundhwes* 2019 Jun;143:22-29. [doi: [10.1016/j.zefq.2019.04.002](https://doi.org/10.1016/j.zefq.2019.04.002)] [Medline: [31204267](https://pubmed.ncbi.nlm.nih.gov/31204267/)]
85. Qiu Y, Yu P. Nursing information systems - applying usability testing to assess the training needs for nursing students. *Methods Inf Med* 2007;46(4):416-419. [doi: [10.1160/me0437](https://doi.org/10.1160/me0437)] [Medline: [17694234](https://pubmed.ncbi.nlm.nih.gov/17694234/)]
86. Nielsen J, Berger J, Gilutz S, Whittenton K. Return on investment (ROI) for usability. Fremont: Nielsen Norman Group; 2003.
87. Welker JA. Implementation of electronic data capture systems: barriers and solutions. *Contemp Clin Trials* 2007 May;28(3):329-336. [doi: [10.1016/j.cct.2007.01.001](https://doi.org/10.1016/j.cct.2007.01.001)] [Medline: [17287151](https://pubmed.ncbi.nlm.nih.gov/17287151/)]
88. Guidance for Industry - Computerized systems used in clinical trials. US Food and Drug Administration. 2018 Mar 11. URL: <https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/fda-bioresearch-monitoring-information/guidance-industry-computerized-systems-used-clinical-trials> [accessed 2021-08-22]
89. Guideline on computerised systems and electronic data in clinical trials. European Medicines Agency. 2021. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/draft-guideline-computerised-systems-electronic-data-clinical-trials_en.pdf [accessed 2021-08-22]
90. Ndlovu K, Mauco KL, Grover S. Evaluation of REDCap as an alternative data management and quality improvement tool for cutaneous lymphoma care in Botswana. 2019 Presented at: 12th Health Informatics in Africa Conference; November 20-22, 2019; Gaborone p. 51-57.
91. Dunn WD, Cobb J, Levey AI, Gutman DA. REDLetr: workflow and tools to support the migration of legacy clinical data capture systems to REDCap. *Int J Med Inform* 2016 Sep;93:103-110 [FREE Full text] [doi: [10.1016/j.ijmedinf.2016.06.015](https://doi.org/10.1016/j.ijmedinf.2016.06.015)] [Medline: [27396629](https://pubmed.ncbi.nlm.nih.gov/27396629/)]
92. Nielsen J, Landauer T. A mathematical model of the finding of usability problems. 1993 Presented at: INTERCHI93: Conference on Human Factors in Computing; April 24-29, 1993; Amsterdam. [doi: [10.1145/169059.169166](https://doi.org/10.1145/169059.169166)]
93. Nielsen J. How many test users in a usability study? Nielsen Norman Group. 2012. URL: <https://www.nngroup.com/articles/how-many-test-users/> [accessed 2021-08-22]
94. Faulkner L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Res Methods Instrument Comput* 2003 Aug;35(3):379-383. [doi: [10.3758/BF03195514](https://doi.org/10.3758/BF03195514)]
95. Hwang W, Salvendy G. Number of people required for usability evaluation. *Commun ACM* 2010 May 01;53(5):130-133. [doi: [10.1145/1735223.1735255](https://doi.org/10.1145/1735223.1735255)]
96. Tenzer F. Anteil der Nutzer von Tablets in Deutschland 2020. Statista. 2020. URL: <https://de.statista.com/statistik/daten/studie/319281/umfrage/anteil-der-tablet-nutzer-in-deutschland/> [accessed 2021-08-22]
97. Etchells E, Bailey C, Biason R, DeSousa S, Fowler L, Johnson K, et al. Human factors in action: getting "pumped" at a nursing usability laboratory. *Healthc Q* 2006;9 Spec No:69-74 [FREE Full text] [doi: [10.12927/hcq.2013.18463](https://doi.org/10.12927/hcq.2013.18463)] [Medline: [17087172](https://pubmed.ncbi.nlm.nih.gov/17087172/)]
98. Aiyegbusi OL. Key methodological considerations for usability testing of electronic patient-reported outcome (ePRO) systems. *Qual Life Res* 2020 Feb;29(2):325-333 [FREE Full text] [doi: [10.1007/s11136-019-02329-z](https://doi.org/10.1007/s11136-019-02329-z)] [Medline: [31691202](https://pubmed.ncbi.nlm.nih.gov/31691202/)]

Abbreviations

digiDEM Bayern: Digital Dementia Registry Bavaria

EDC: electronic data capture

REDCap: Research Electronic Data Capture

SUS: System Usability Scale

TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 29.06.21; peer-reviewed by S Santini, M Broekhuis; comments to author 22.07.21; revised version received 23.08.21; accepted 19.09.21; published 03.11.21.

Please cite as:

Reichold M, Heß M, Kolominsky-Rabas P, Gräbel E, Prokosch HU

Usability Evaluation of an Offline Electronic Data Capture App in a Prospective Multicenter Dementia Registry (digiDEM Bayern): Mixed Method Study

JMIR Form Res 2021;5(11):e31649

URL: <https://formative.jmir.org/2021/11/e31649>

doi: [10.2196/31649](https://doi.org/10.2196/31649)

PMID: [34730543](https://pubmed.ncbi.nlm.nih.gov/34730543/)

©Michael Reichold, Miriam Heß, Peter Kolominsky-Rabas, Elmar Gräbel, Hans-Ulrich Prokosch. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.11.2021. 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

Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland

Kristjana Thorarinsdottir¹, MSc; Emily A Holmes^{2,3}, Prof Dr; Johann Hardarson¹, MSc; Unnur Hedinsdottir¹, MSc; Marie Kanstrup³, PhD; Laura Singh^{2,4}, PhD; Arna Hauksdottir⁵, Prof Dr; Thorhildur Halldorsdottir^{1,5}, PhD; Berglind Gudmundsdottir^{6,7}, Prof Dr; Unnur Valdimarsdottir^{5,8,9}, Prof Dr; Edda Bjork Thordardottir⁵, PhD; Beau Gamble², PhD; Andri Bjornsson¹, Prof Dr

¹Department of Psychology, University of Iceland, Reykjavik, Iceland

²Department of Psychology, Uppsala University, Uppsala, Sweden

³Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

⁴Swedish Collegium for Advanced Study, Uppsala, Sweden

⁵Centre of Public Health Sciences, University of Iceland, Reykjavik, Iceland

⁶Faculty of Medicine, University of Iceland, Reykjavik, Iceland

⁷Landspítali - The National University Hospital of Iceland, Reykjavik, Iceland

⁸Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

⁹Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States

Corresponding Author:

Kristjana Thorarinsdottir, MSc

Department of Psychology

University of Iceland

Sæmundargata 12

Reykjavik, 101

Iceland

Phone: 354 6942001

Fax: 354 6942001

Email: kth35@hi.is

Related Article:

This is a corrected version. See correction statement: <https://formative.jmir.org/2021/11/e34897/>

Abstract

Background: Additional interventions are needed for survivors of psychological trauma because of several barriers to and limitations of existing treatment options (eg, need to talk about the trauma in detail). Case studies are an important step in exploring the development of novel interventions, allowing detailed examination of individual responses to treatment over time. Here, we present a case study that aims to test a novel intervention designed to disrupt memory reconsolidation, taking a single-symptom approach by focusing on intrusive memories of a traumatic event.

Objective: This study aims to examine a novel brief cognitive intervention to reduce the number of intrusive memories of trauma in an Icelandic setting and to extend previous studies by examining long-term effects for up to 3 months. The intervention was guided by a clinical psychologist and comprised a brief memory reminder, followed by Tetris gameplay with mental rotation, targeting one memory at a time in each session.

Methods: This was a single case study in Iceland with a woman in her 50s (drawn from an epidemiological study of trauma) with subthreshold posttraumatic stress disorder and a diagnosis of obsessive-compulsive disorder and social anxiety disorder. The participant had four different intrusive memories from a traumatic event that happened in her childhood. The primary outcome was the change in the number of intrusive memories from baseline to intervention phase and to follow-ups. The number of intrusions was monitored in a daily diary for 4 weeks preintervention, 8 weeks during the intervention, and 1 week at 1-month and 3-month follow-ups. Intrusions were targeted one by one over six intervention sessions, creating four *repetitions* of an *AB design* (ie, length of baseline *A* and intervention phase *B* varied for each memory). We examined changes in both *the total* number

of intrusions (summed across all four memories) and individually for each memory. In addition, we explored whether having fewer intrusive memories would have an impact on functioning, posttraumatic stress, and depression or anxiety symptoms.

Results: The total number of intrusions per week was 12.6 at baseline, 6.1 at the intervention phase (52% reduction from baseline), 3.0 at the 1-month follow-up (76% reduction), and 1.0 at the 3-month follow-up (92% reduction). Reductions in the symptoms of posttraumatic stress and depression were observed postintervention. Sleep, concentration, stress, and functioning improved. The participant considered the gameplay intervention acceptable and helpful in that she found that the memories disappeared while she was playing.

Conclusions: This guided brief cognitive intervention reduced the number of intrusive memories over the intervention phase and follow-ups. The brief memory reminder was well tolerated, removing the need to discuss trauma in detail. The next steps require an extension to more cases and exploring remote delivery of the intervention.

(*JMIR Form Res* 2021;5(11):e29873) doi:[10.2196/29873](https://doi.org/10.2196/29873)

KEYWORDS

psychological trauma; intrusive memories; case report; visuospatial interference task; Tetris gameplay; mental imagery; mobile phone

Introduction

Background

Psychological trauma (eg, disasters, accidents, or interpersonal violence) is experienced by most people at some point during their lifetime [1,2]. Many individuals who have been exposed to trauma (approximately 1 in 4) go on to develop posttraumatic stress disorder (PTSD) [3,4]. The core clinical symptom of PTSD is intrusive memories related to traumatic events [1,5]. Other symptoms of PTSD include avoidance of stimuli associated with trauma, along with negative alterations in cognition, mood, arousal, and reactivity [1]. Approximately half of those diagnosed with PTSD do not spontaneously recover within 40 months of diagnosis [6]. PTSD, even when subthreshold, is associated with substantial distress, functional impairment, and comorbidities [1,7]. Although many patients respond well to current PTSD treatments, approximately one-third of patients who enter psychological treatment for PTSD still meet the diagnostic criteria for the disorder following treatment [8].

Current evidence-based treatments for PTSD include individualized trauma-focused cognitive behavioral therapy interventions and eye movement desensitization and reprocessing [3,8]. However, there are some limitations to existing psychological treatment options for PTSD, including the limited number of qualified therapists, geographic distances to such clinical expertise (eg, in rural areas), high cost of treatment, and stigma being a barrier to individuals seeking treatment [9]. Dropout rates from PTSD treatment are high, approximately 18% overall (ranging from 0%-48%) in clinical trials and are thought to be even higher in clinical practice outside of clinical trials [10-12]. Furthermore, only a minority of those who need PTSD treatment receive it [13]. The common denominator in existing treatment options is a requirement for patients to recall and talk about the traumatic experience in detail, which many trauma survivors are reluctant to do [9]. Many therapists are also reluctant to deliver trauma-focused therapies, such as prolonged exposure, because of fear of exacerbation of symptoms or concerns with patient dropout [12].

Another barrier to treatment is the lack of service provision [9]. Iceland, for instance, is one of many countries that lack the mental health services capacity to offer treatment to all trauma survivors. New, briefer approaches that reach more people or can be delivered to remote places in geographically dispersed countries via the internet are needed [9]. Moreover, people who do not meet the full diagnostic criteria for PTSD are typically unable to access existing services, meaning that treatments for trauma survivors with subthreshold but impairing symptoms are needed.

Overall, these limitations and barriers create the need for additional complementary approaches to current treatments. One option that has been suggested is to focus on reducing one single, tractable symptom (here, the core clinical symptom) rather than a full diagnosis of PTSD [5,14]. Intrusive memories (ie, criterion B1 as defined in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition [DSM-5]*) are repeated and unwanted memories of scenes from a traumatic event, and they are predominantly visual [15-17]. They can evoke the same emotions experienced during the traumatic event [16] and often have a sense ofnowness, that is, as if they are happening in the present rather than in the past [15]. Intrusive memories can cause significant distress and interfere with everyday functioning, making them an important target for treatment [17].

A relatively simple and brief intervention to reduce the number of intrusive memories after trauma has recently been developed, building on principles from cognitive science [14,18,19]. It is in line with calls to develop new therapeutic approaches for PTSD, such as those that target memory reconsolidation [20]. The intervention comprised a brief memory reminder for a specific intrusive memory of trauma, practice in mental rotation (ie, actively playing the game by rotating the blocks in one's mind; for further details, see Holmes et al [21], chapter 11), followed by Tetris gameplay with mental rotation for 25 minutes, guided in person by a researcher.

Initial work toward clinical translation was for recent memories of trauma [22-24]. For older, intrusive memories of trauma, the effect has been explored using case study and case series approaches [19,25,26]. Kessler et al [19] conducted a case series

of inpatients ($n=20$) with complex PTSD and trauma memories from childhood. The intervention comprised a memory reminder (here, writing a brief description of the memory, then shredding it) followed by Tetris gameplay with mental rotation for 25 minutes for one intrusive memory at a time. Memories (here, many different memories) were targeted one by one, that is, each intrusive memory in a different session, and memories were tracked individually in a diary. The results showed that targeting a specific intrusion was followed by a drop in the frequency of that intrusion (some to zero). The frequency of targeted intrusions reduced by 64% overall from baseline to postintervention, whereas the frequency of nontargeted intrusive memories reduced by 11%.

Kanstrup et al [25] adapted the intervention for a new target group—people who were refugees ($n=4$) and used it to target already established trauma memories such as of war. The memory reminder used here was a brief list of intrusive memories (ie, hotspot sheet) where participants were asked to briefly describe in a few words the imagery content of their intrusions, either by writing it themselves or by telling the researcher what to write. The intervention was delivered in a community setting, such as a library. All 4 participants showed a decrease in the number of intrusive memories (again targeted one by one) after the intervention and reported improved functioning. For example, participant 1 had a decrease from 10 memories at baseline to 0 after the first intervention week, and for participant 3, the 28 memories at baseline decreased to 14 after the first intervention week.

Study Design and Aims

Given the small-scale but promising results of this single-symptom intervention approach for older memories of trauma, we were interested in adapting it for women with a trauma history in Iceland. Thus, in this case study ($n=1$), we aim to investigate the effects of the intervention adapted to Iceland for a woman from a population-based sample experiencing intrusive memories of childhood trauma, delivered with guidance from a clinical psychologist and seen in a university research setting. Importantly, we aim to extend the previous literature testing this intervention by including a significantly longer follow-up period than previous studies (ie, 1 month and 3 months postintervention) to examine whether effects are maintained in the long term.

When evaluating the efficacy of novel interventions and refining intervention protocols, single-case designs are a crucial step [27,28], giving researchers the chance to examine individual variability over time [29]. The *N-of-1* trials are also gaining popularity as modern medicine moves toward individualized patient-centered care [30]. Typically, a replication of AB (ABAB) is considered necessary to establish intervention effects [27]. However, Kanstrup et al [25] argued that a classic ABAB design was not optimal for evaluating this specific intervention as, unexpectedly, the effects lasted after one intervention session and did not rebound (ie, could not be reduced again, as assumed

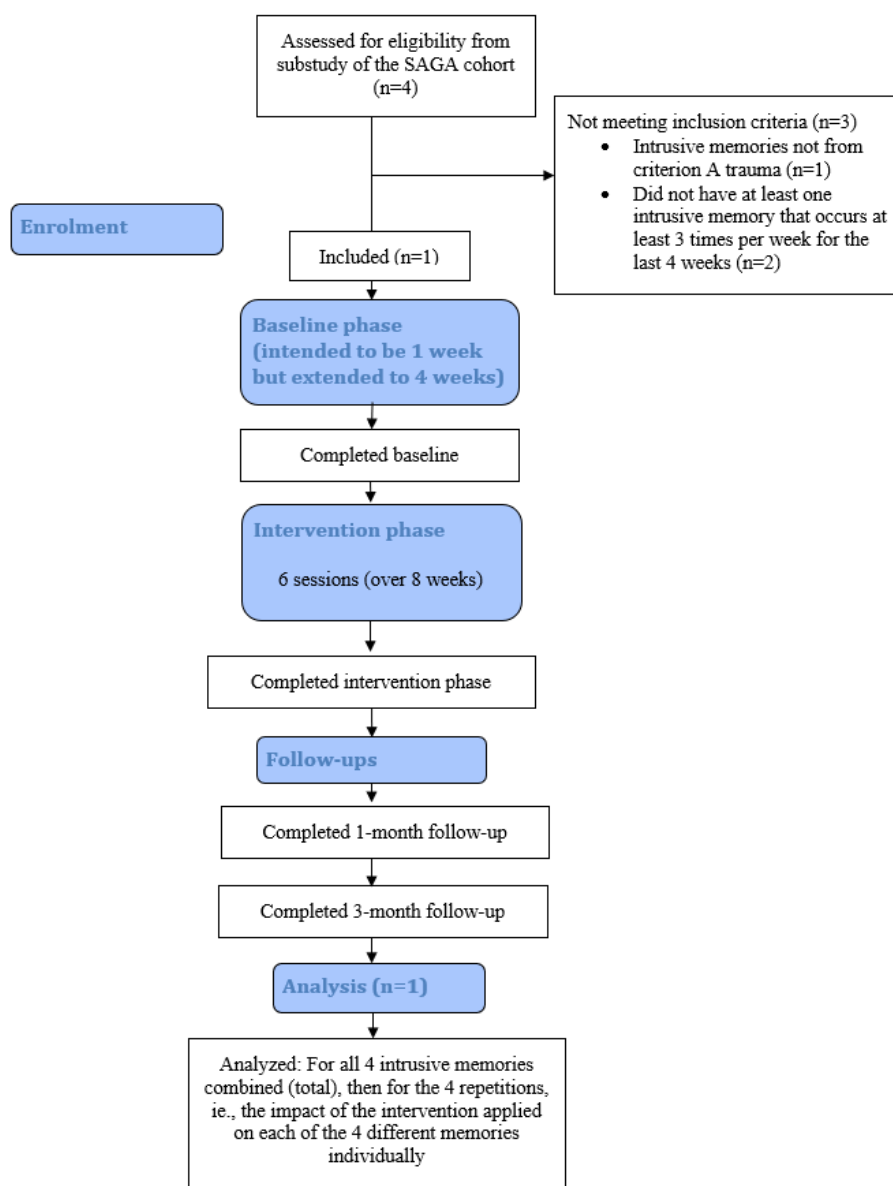
by a classic ABAB). Kanstrup et al [25] instead recommended a within-person *multiple baseline AB design*, as in the study by Kessler et al [19]. In this approach, if a person has more than one different intrusive memory, then each specific intrusive memory is targeted one at a time, with separate intervention sessions allowing focused assessment of the effect of each intervention on each memory over time. This is the design adopted in this study. However, to avoid confusion with other case series designs (such as those with multiple *randomized baselines*), we refer to this design as a *repeated AB design*.

We predict that our participant (here with four different intrusive memories of trauma) would report fewer intrusive memories (primary outcome) during the intervention phase than in the preceding baseline phase and that the reduction in the number of intrusions would be maintained at the 1-month and 3-month follow-ups. We also aim to explore whether having fewer intrusive memories would be associated with improvements in general functioning and reductions in symptoms of PTSD, depression, and anxiety (secondary outcomes). In addition, we aim to explore the feasibility and acceptability of the intervention (similar to Holmes et al [31]).

Methods

Participants

Women who took part in a substudy of the stress-and-gene-analysis (SAGA) cohort study were screened for eligibility. The SAGA cohort study is a population-based longitudinal cohort study of Icelandic women who completed an extensive questionnaire on trauma history and mental health (baseline data collection finished on July 1, 2019). The substudy (the Social Trauma Project) involves comparing two samples of women from the SAGA cohort study with either likely PTSD (ie, having a score on the PTSD Checklist-5 [PCL-5; see the *Measures* section] of ≥ 33) or not likely PTSD (ie, scores in the lowest one-fifth on the PCL-5), using clinical interviews. When taking part in the substudy, two semistructured interviews were administered (ie, the Mini International Neuropsychiatric Interview [MINI], also used to assess the exclusion criteria for this study, and the Clinician Administered PTSD Scale [CAPS]; see the *Measures* section). When taking part in the substudy, women were screened for the presence of intrusive memories of trauma. The screening included a short description of the symptom, followed by questions about the presence of the symptom to assess their eligibility for this study. A total of 4 women from the substudy who provided consent to be contacted regarding additional research were assessed for inclusion in this case study. A total of 3 women did not meet the inclusion criteria (CONSORT [Consolidated Standards of Reporting Trials] diagram in Figure 1). The included participant was a woman in her 50s who had four different intrusive memories from a single traumatic event involving physical violence in childhood (ie, occurring around four decades previously).

Figure 1. Adapted Consolidated Standards of Reporting Trials participant flow diagram for this study. SAGA: Stress and Gene Analysis.

The inclusion criteria were as follows: (1) having experienced criterion A trauma as defined by *DSM-5* [6]; (2) having at least one intrusive memory that occurs at least three times per week for the last 4 weeks; (3) being able and willing to attend three to eight sessions with the researcher; (4) being able and willing to monitor intrusive memories in daily life; (5) having access to a smartphone; and (6) being able to speak Icelandic and read study materials in Icelandic. Exclusion criteria were as follows: (1) current psychotic disorder; (2) current manic episode; and (3) being acutely suicidal. Exclusion criteria were assessed with the MINI.

The participant reported clinically significant past-month PTSD symptoms from physical violence experienced in childhood, with a total symptom severity score of 22 of 80 on the CAPS and missing one symptom in the E cluster to meet full diagnostic criteria (had five symptoms in cluster B, two in cluster C, three in cluster D, and one in cluster E). This assessment took place 2 months before participation in this study as part of the substudy of the SAGA cohort using the CAPS (see the *Measures*

section). The participant also met criteria for social anxiety disorder and obsessive-compulsive disorder according to the MINI diagnostic interview (see the *Measures* section). The participant received psychological treatment in the past for problems related to work but had never received trauma-related psychological treatment. She reported not taking any psychotropic medication in the 3 months before taking part.

Design

This single case study took a specific single-symptom probe approach, whereby each of the four intrusive memories was targeted one at a time in different sessions [19]. Critical to this approach, the participant distinguishes the content of their different intrusive memories (here, for four intrusions, eg, (1) red curtain, (2) man's face, (3) blood on floor, and (4) closed door; these examples are fictitious to protect anonymity) and tracks the frequency of each intrusion over time. We describe this design here as a *repeated AB design*, wherein the length of baseline (A, preintervention; monitoring only) and intervention (B) phases varied across each of the four intrusive memories,

depending on when each memory was targeted. The baseline phases for each individual memory are used as control periods to compare their numbers before and after being targeted by the intervention.

The number of each intrusive memory was monitored in a daily diary for 4 weeks preintervention, over 8 weeks of the intervention, and then for 1 week at the 1-month and 3-month follow-ups, that is, the participant monitored the occurrence of her intrusive memories in a daily diary before each intervention session to establish a baseline level of intrusion. This baseline phase was intended to be 1 week; however, the diary was kept for 4 weeks, as the participant was not able to meet with the researcher when planned. The intervention phase lasted 8 weeks rather than 6 weeks, as planned for the same reason. However, the participant did monitor the frequency of her intrusive memories in these extended periods, and we included all the data in the analyses. The daily diary was kept again for 1 week at the 1-month and 3-month follow-ups.

The participant's four different intrusive memories were targeted one by one over six intervention sessions guided by a clinical psychologist who specialized in trauma-focused cognitive behavioral therapy. The design thus involved four repetitions of an AB design. In addition to the six guided sessions, the participant could also self-administer the intervention at any time after the first session if she so chose for memories already targeted in the session. The primary outcome was the change in the number of intrusive memories from baseline to the intervention phase and to long-term follow-ups (1 and 3 months). The participant also completed self-report measures for PTSD, depression and anxiety symptoms, and functional impairment at baseline, the last intervention session, and the 1-month and 3-month follow-ups.

Procedure

Training

To promote adequate intervention delivery and protocol adherence, the researcher delivering the intervention (JPH, a licensed clinical psychologist and specialist in trauma therapy) received training and clinical supervision from experienced researchers or clinical psychologists who had expertise in delivering the novel intervention (EAH and MK). Training included two in vivo workshops for 3 days and then approximately 6 months later for 2 days. Workshops covered theoretical and practical aspects of intervention delivery and included role-plays with trainers until adequate performance was reached. Training also included how to explain and capture the primary outcome measure (intrusive memory diary). During data collection, the researcher received continued supervision, adherence checks, and support regarding any adaptations necessary from a clinical supervisor via telephone after sessions with the participant and weekly supervision meetings. The researcher also participated in remote group training meetings twice a month with other researchers using the intervention.

Baseline Session

In the first session, the participant answered baseline questionnaires (relating to secondary outcomes), and the researcher explained what intrusive memories are (ie, memories

that include sensory impressions such as sight and sound; are predominantly visual in form, similar to pictures or a film clip in the mind's eye; and are distressing and occur involuntarily). The participant identified her different intrusive memories by briefly describing them to the researcher using only a few words to indicate their visual content; the researcher wrote the description on a *hotspots* sheet that was clearly visible to the participant. The participant did not talk about the trauma with the researcher or about the intrusive memories in detail. The participant labeled each of her intrusive memories with a symbol (ie, first memory labeled *A* and second memory labeled *B*) and was instructed on how to monitor the daily frequency of them in a diary (primary outcome measure). When indicating experiencing a memory, the participant noted the symbol corresponding to that specific memory in a specific time frame of that day. Each diary included 7 days and four periods each day (see the *Measures* section).

Intervention Sessions

In each of the intervention sessions (six sessions), the participant selected one memory at a time to target that week and completed the intervention procedure, guided by the clinical psychologist. The intrusion selected first can be the one that is most troublesome or frequent or one for any other reason the participant wishes to try reducing first. The intervention consisted of a brief memory reminder (ie, briefly thinking about the intrusive memory to bring the image to mind without it becoming emotionally overwhelming; this approach is different from the memory reminder used by Kessler et al [19]). After the memory reminder, the participant was trained in mental rotation, followed by Tetris gameplay for 25 minutes with an emphasis on mental rotation (see Holmes et al [21], chapter 11). The Tetris gameplay was delivered with the videogame Tetris DS in the Nintendo DS, set to *marathon* mode and ghost piece off, on a 10.1-inch screen. Between sessions, the participant was invited to self-administer the intervention using a Tetris app [32] on her smartphone, that is, to repeat the intervention for already targeted intrusions (instructed to play in the same way as in session when the intrusion came to mind involuntarily). Only one intrusion was targeted per session; when the next intrusion was targeted, the participant again (not the therapist) selected the memory to target. At the start of the last intervention session, the participant also completed the secondary outcome measures.

Follow-up

At the 1-month and 3-month follow-ups, the participant recorded the number of intrusions in the diary daily for 1 week and completed secondary outcomes. All data were recorded on a laptop computer using the REDCap (Research Electronic Data Capture) database, an encrypted electronic software, and stored on secure servers [33]. At the 1-month follow-up, the participant was in quarantine because of the COVID-19 pandemic, and thus, all follow-up measures were administered remotely through the REDCap platform; see the *Procedure* section.

Measures

Eligibility Assessments (Part of the SAGA Cohort Substudy)

The CAPS-5 is a 30-item semistructured interview used to assess symptoms of PTSD from physical violence in childhood and symptom severity in the past month, according to the *DSM-5* [1]. Each item is scored on a 5-point Likert scale (0=mild or subthreshold; 4=extreme or incapacitating) with a threshold symptom rating of 2 (ie, moderate) for a possible diagnosis. Frequency and intensity of each symptom were assessed and rated separately. The CAPS-5 has excellent internal consistency (Cronbach $\alpha=.88$) and test-retest reliability (0.83), along with good convergent validity (0.83 [34]), making it a useful tool for diagnosing PTSD.

The MINI is a structured diagnostic interview that assesses axis 1 psychiatric disorders according to the *DSM-4*. The MINI has been shown to have good sensitivity and specificity for most diagnoses [35]. Interrater and test-retest reliability has been shown to be good, with kappa values in the high to very high range ($\kappa=0.79-1.00$ [36]).

Primary Outcome Measure

The intrusive memory diary was adapted from previous studies [22,25]. Each diary included a daily pen-and-paper record of four timeframes per day (morning, afternoon, evening, and night) for 7 days. Instructions on how to use the diary included a definition of intrusive memories of trauma as mental images (in the form of pictures or a film clip in the mind's eye) that are distressing and occur involuntarily. The participant was instructed not to record voluntary thoughts or verbal thoughts about the trauma without sensory content. The participant monitored the occurrence of her intrusive memories in a daily diary for 4 weeks before any intervention sessions, for 8 weeks while intervention sessions were administered, and again for 1 week at the 1-month and 3-month follow-ups. Throughout this, the participant noted which of the four different memories each intrusion was, allowing us to examine changes in each memory individually. The primary outcome was the change in the number of intrusive memories from baseline to the intervention phase and to long-term follow-ups (1- and 3-month follow-ups).

Secondary Outcome Measures

PTSD symptoms were assessed with the PCL-5, a 20-item self-report scale used to assess the severity of PTSD symptoms in the past month from physical violence in childhood, corresponding to the *DSM-5* criteria for PTSD [34]. Each symptom is rated on a 4-point Likert scale (0=not at all; 4=extremely). The PCL-5 has strong internal and test-retest reliability, with good convergent and discriminant validity [37]. The Icelandic translation of the PCL-5 had excellent internal consistency in the SAGA cohort study ($\alpha=.95$). Assessment of clinical significance is not yet clear for the PCL-5; however, a score of 33 is likely to correspond to a *DSM-5* PTSD diagnosis, and a score of ≤ 24 posttreatment is likely to represent clinically significant change [38].

Depression symptoms were assessed with the Patient Health Questionnaire-9 (PHQ-9), a nine-item self-report measure of

depressive symptoms and their severity in the prior 2 weeks [39]. Each item is rated on a 4-point Likert scale (0=not at all; 3=nearly every day). The PHQ-9 has excellent internal reliability (Cronbach α ranging from .86 to .89) and good test-retest reliability ($r=0.84$ [39]). The Icelandic version had good internal consistency in the SAGA cohort study ($\alpha=.89$). A five-point change in the PHQ-9 score is considered clinically significant [40].

Anxiety symptoms were assessed with the Generalized Anxiety Disorder-7 (GAD-7) scale, a brief self-report questionnaire used as a screening tool for GAD symptoms and their severity in the prior 2 weeks [41]. Each item is rated on a 4-point Likert scale (0=not at all; 3=nearly every day). The GAD-7 has excellent internal consistency (Cronbach $\alpha=.92$) and good test-retest reliability ($r=0.83$ [41]). The GAD-7 has been reported to be useful in screening for anxiety disorders in general [42]. The Icelandic version had good internal consistency in the SAGA cohort study ($\alpha=.90$). A four-point change in the total score is considered clinically significant on the GAD-7 [43].

Functional impairment was assessed with the Sheehan Disability Scale (SDS), a self-report measure designed to assess functional impairment in the prior week across three domains: (1) work or school, (2) social, and (3) family life [44]. These domains are measured on an 11-point scale (0=not at all; 10=extremely). The scale was adjusted to assess functional impairment associated with intrusive memories. This scale has been shown to have good psychometric properties [44]. A three-point change in the SDS score has been used as a measure of treatment response [45]. The Icelandic version has good internal consistency in clinical groups ($\alpha=.70-.84$ [46]).

Self-guided adherence to the use of the gameplay intervention in daily life was assessed with a question regarding how often Tetris was played after experiencing an intrusive memory (11-point scale; 0=not at all; 10=every time).

Feasibility and acceptability rating for using the smartphone gameplay intervention was assessed with two self-rated items: whether the participant would recommend the intervention to a friend and whether she thought gameplay was an acceptable way to reduce intrusive memories. Scores could range from 0 to 10, with higher scores indicating greater acceptability or feasibility. Two open-ended questions were also asked: "How did you feel about playing Tetris after you had an intrusive memory?" and "Did you find the intervention helpful? If yes, how?"

The impact of intrusive memories on concentration, sleep, and stress was assessed with six self-rated items about the past week: two items assessing concentration difficulties in general and because of intrusive memories (11-point scale; high scores indicating more difficulties); one item assessing duration of disruption after experiencing intrusive memories (five response options ranging from <1 minute to >60 minutes); two items assessing sleep disturbances because of intrusive memories (sleep in general and nightmares; 11-point scale; higher scores indicating more sleep disturbance); and one item assessing the degree to which intrusive memories affected stress levels (0=not at all; 10=affected very much).

Ratings of the general impact of intrusive memories were obtained with two items: one assessing distress caused by intrusive memories and the other assessing how vivid they were in the past week, both rated on an 11-point scale (0=not at all; 10=very distressing or vivid).

Intrusion diary adherence was assessed with one item addressing the accuracy of filling out the diary (0=not at all; 10=very accurately).

The impact of intrusive memories on daily functioning was assessed with two items. One question was open-ended: "How have the intrusive memories affected your ability to function in your daily life in the past week?" The other question was self-rated: "Have the intrusive memories affected your ability to function in your daily life?" (11-point scale, a higher score indicating a greater impact on functioning).

Data Analysis

Changes in the Total Number of Intrusive Memories

The primary outcome was change in the number of intrusive memories from baseline to the intervention phase and to long-term follow-ups (1 month and 3 months). We first examined the primary outcome in terms of the *total* number of intrusions (before examining separately for each memory). For this, we summed the number of all intrusions occurring across the 4-week baseline period, then across the 8-week intervention period, and then at each of the 1- and 3-month follow-ups. Given that these periods differed in duration, we calculated the total number of intrusions *per week* to generate a measure that was comparable across periods. Missing data were dealt with by excluding these time points from calculations. For example, the baseline period was 29 days, but data were present for 22.25 days; thus, the total number of intrusions per week was calculated as $40 \text{ intrusions} / 22.25 \text{ days} \times 7 = 12.6 \text{ intrusions per week}$ at baseline.

To examine changes over time, we calculated the percentage reduction in total intrusions per week from baseline to the other periods. For example, as there were 6.1 intrusions per week in the intervention phase, this was calculated as $(1 - [6.1 / 12.6]) \times 100 = 52\%$ reduction in the intervention phase compared with baseline.

Change in the Number of Each of the Four Specific Intrusive Memories

Next, we examined the data per intrusive memory. Here, each intrusion acts as its own control, that is, the specific baseline

phase for each individual memory is used as a control period to compare its number before and after being targeted by the intervention. There is a different baseline (*A*) and intervention (*B*) phase per memory, depending on which session it was targeted. The percentage reduction in each intrusion after being targeted was calculated as $1 - (\text{mean number per week during intervention phase} / \text{mean number per week during baseline}) \times 100$. Percentage reductions were then calculated in the same way for the 1- and 3-month follow-ups compared with baseline.

Other Symptoms and Functioning

We also used a descriptive approach to investigate whether there were clinically significant changes over time in the overall symptoms of PTSD, depression, anxiety, and functional impairment.

Ethics Statement

The study was approved by the National Bioethics Committee of Iceland (Number VSNb2017110046/03.01). The participant provided written and informed consent. All sessions followed a written protocol. No adverse events were reported by the participant.

Open Science Statement

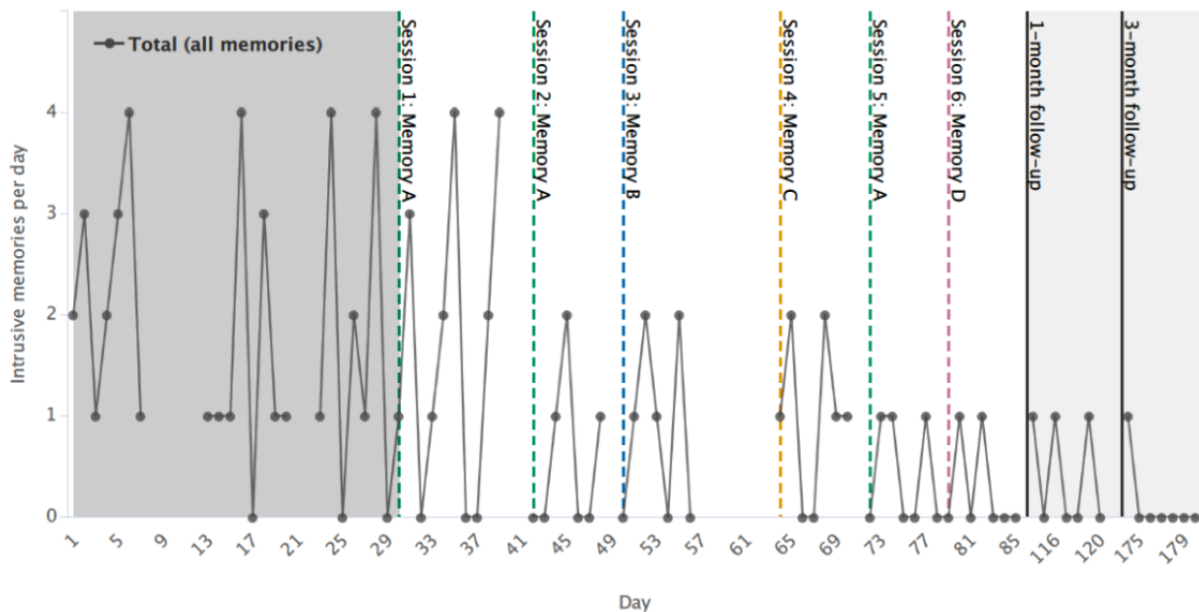
This single case study was not preregistered but precedes and is similar to the design and procedures of a case series ($n=5$) that we later registered on ClinicalTrials.gov (NCT04209283) on December 4, 2019. All anonymized summary-level data are reported in this manuscript. Study materials may be made available upon reasonable request with an appropriate materials transfer agreement with University of Iceland. It should be noted that the delivery of this intervention requires extensive training and supervision (see the Procedure: Training section).

Results

Overview

The participant had four different intrusive memories that were all predominantly visual and tracked each intrusion over time. All her intrusive memories were from a single traumatic event that took place roughly four decades before participation. All four intrusive memories were targeted with the intervention at different time points during the intervention phase (Figure 2).

Figure 2. Graph for visual inspection of primary outcome data (total number of intrusive memories) on the y-axis as total per day. Days since enrollment is shown on the x axis, which includes baseline (gray), intervention (white), and follow-up periods (light gray). Dashed colored vertical lines show when each intervention session was administered and which specific memory (memories A, B, C, or D) was targeted (eg, session 1; memory A in green). Memories are labeled in the order of when they were targeted (eg, memory A was targeted in the first intervention session). Solid black vertical lines show the 1-month and 3-month follow-ups. Gaps in the time series in the baseline and intervention periods reflect the missing data (for each specific intrusive memory data, see Figure 3).



One memory (memory A) was targeted three times (reported as the most distressing and frequent by the participant), and the other memories were targeted once. The participant readily understood the instructions given and successfully completed the intervention sessions and procedures. Intrusive memory diary data were missing for days 8-12 and 21-22 during the baseline phase and for days 40-41, day 49, days 57-63, and day 71 during the intervention phase; the diary was fully completed at follow-ups. Most missing diary data were because of extra days passing in between sessions, that is, when the participant had completed their current diary (covering a period of only 1 week) but had not received their next diary. No attempt was made to retrieve data for the missing days. In total, the diary

was completed successfully for 82% (81/99 days) of the study period.

Primary Outcome

Change in the Total Number of Intrusive Memories

Across the 4-week baseline period, the total number of intrusions was relatively stable and approximately 12.6 per week (summed across all four memories). This number reduced to 6.1 per week across the 8-week intervention phase (52% reduction from baseline), to 3.0 per week at 1-month follow-up (76% reduction), and to 1.0 per week at the 3-month follow-up (92% reduction; Table 1).

Table 1. Number of intrusive memories per week at baseline, intervention, 1-month follow-up, and 3-month follow-up, and relative reduction (in percentage) from baseline for total intrusions and for each memory separately (n=1).

Intrusions	Baseline (A; number per week)	Intervention (B; number per week)	Reduction (%)	1-month follow-up (number per week)	Reduction (%)	3-month follow-up (number per week)	Reduction (%)
Total ^a	12.6	6.10	52	3.0	76	1.0	92
Memory A	3.8	2.0	46	1.0	74	1.0	74
Memory B	3.6	0.5	86	0	100	0	100
Memory C	2.4	1.0	59	1.0	59	0	100
Memory D	1.4	1.0	26	1.0	28	0	100

^aTotal intrusions are not equal to the sum of the intrusions for each memory. This is because the length of the baseline and intervention phases differ across memory and the total. See the Data Analysis section for more details on how these numbers were calculated.

Figure 2 displays the total number of intrusive memories per day (summed across all four intrusive memories) throughout all phases. Visual inspection indicated that after the second intervention session, the total number of intrusions reduced. The number of intrusive memories remained relatively stable

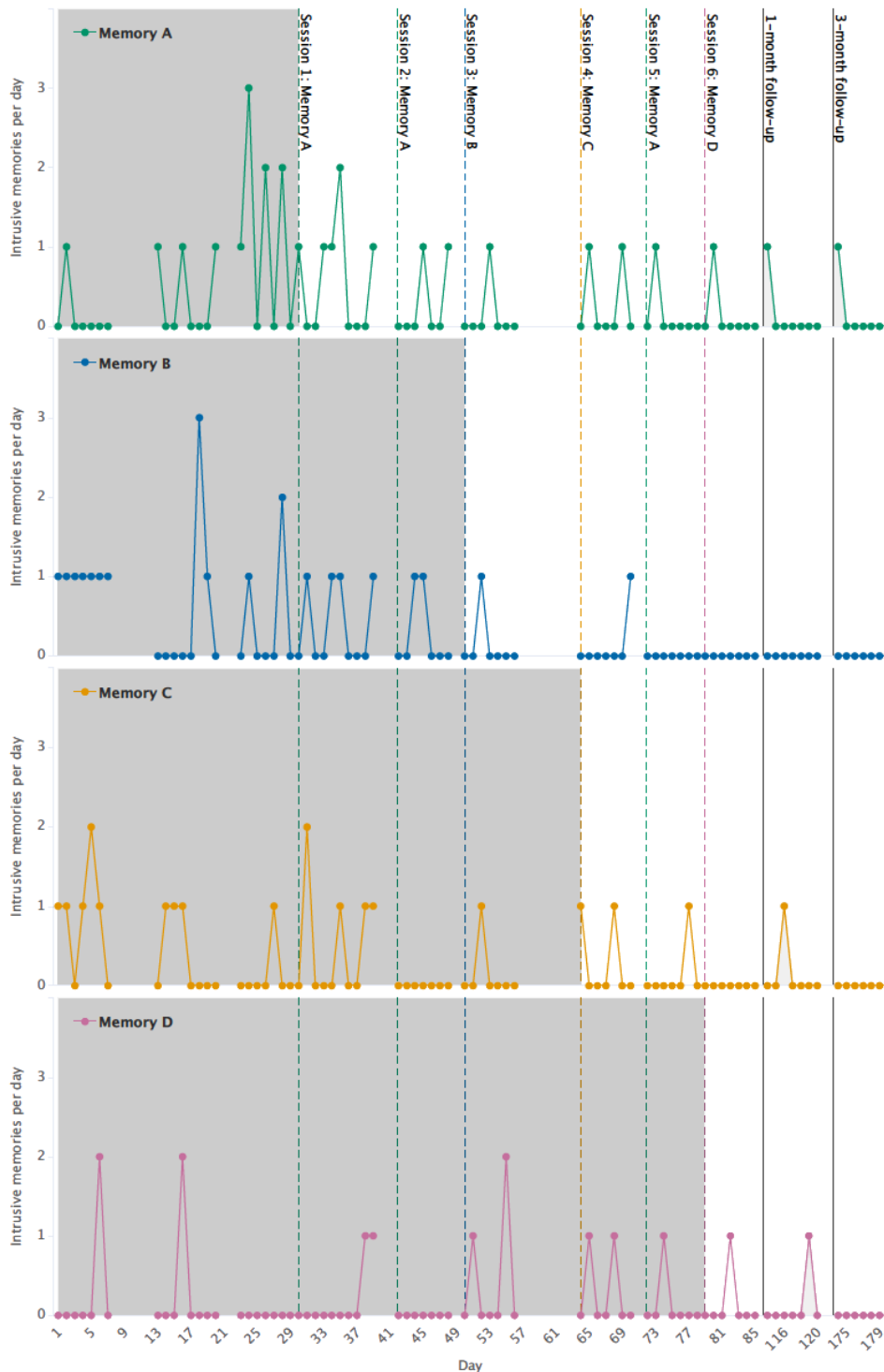
between sessions 2 (day 43) and 5 (day 73) when a further drop in frequency was evident, maintained at the 1-month follow-up, and then continued to drop further at the 3-month follow-up.

Change in the Number of Each of the Four Specific Intrusive Memories

Figure 3 displays the frequency of each intrusive memory during all phases (baseline, intervention, and 1-month and 3-month follow-ups). All four intrusive memories dropped in number

per week after being targeted, that is, reductions of 46%, 86%, 58%, and 26% for memory A, B, C, and D, respectively, from their specific baselines to intervention periods. Three of the four intrusions were eliminated completely at the 3-month follow-up (Table 1).

Figure 3. Graph for visual inspection of the number of intrusive memories (on the y-axis as number per day) for each of the four specific intrusive memories reported by the participant (memories A, B, C, and D). Days since enrollment is shown on the x-axis, which includes baseline (gray), intervention (white), and follow-up periods (light gray). Different baseline and intervention lengths for each memory reflect that this is a repeated AB design. Dashed colored vertical lines show when each intervention session was administered and which specific memory was targeted (eg, session 1: memory A in green). Memories are labeled in the order of when they were targeted (eg, memory A was targeted in the first intervention session). Solid black vertical lines show the 1-month and 3-month follow-ups. Gaps in the time series in the baseline and intervention periods reflect missing data.



Memory A was targeted in intervention sessions 1, 2, and 5 at the participant's request. Visual inspection of Figure 3 shows a drop in frequency in the week after intervention session 1 and a further decrease in the week after session 2. The reduction appears to be stable at the 1-month and 3-month follow-up. However, there appears to be an increase in frequency between days 65 and 73, which resulted in that intrusion being targeted again. In session 5, the participant disclosed that she had come across a person who was present during the traumatic experience (ie, seeing the person triggered that memory).

Memory B was targeted in session 3, and a drop in frequency was evident in the subsequent week, which was maintained throughout the follow-ups. A drop in frequency for memory C was shown in the week after intervention session 1 (memory A targeted), and the frequency reduction remained stable at follow-up. Less changes in frequency were visible for memory

D throughout the intervention phase (targeted in intervention session 6), whereas there was a reduction in frequency at the 3-month follow-up.

Secondary Outcomes

Ratings of Adherence and General Impact of Intrusive Memories

Table 2 shows that the participant rated her intrusions in general as becoming less vivid and distressing over the intervention and follow-up phases. Ratings of self-guided adherence to Tetris gameplay between sessions are also shown in Table 2, indicating that it was most used during the intervention period, and self-reported accuracy for completing the intrusive memory diary was high throughout the study period (mean 8.25, SD 0.5).

Table 2. Ratings of adherence to intrusive memory diary and general impact of intrusive memories (n=1).

Item	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	1-month follow-up	3-month follow-up
Diary accuracy ^a	8	9	8	9	8	9	7	8
Intrusions vividness ^b	7	8	6	9	8	6	4	3
Intrusions distress ^c	6	6	4	8	5	4	3	1
Tetris gameplay ^d	N/A ^e	4	10	5	2	4	1	0

^aHow accurately did you fill out the diary? 0=not at all; 10=very accurately.

^bDuring the last week, how vivid were your intrusive memories? 0=not at all; 10=very vivid.

^cDuring the last week, how distressing were your intrusive memories? 0=not at all; 10=very distressing.

^dHow often did you manage to play Tetris after you experienced an intrusive memory? 0=never; 10=every time.

^eN/A: not applicable.

Feasibility and Acceptability for Using a Smartphone Gameplay Intervention

The participant rated whether she would recommend the intervention to a friend as 10/10 (meaning she would certainly recommend it). She also rated whether she considered gameplay to be an acceptable way to reduce intrusive memories as 10/10 (very acceptable). When asked how she felt about playing Tetris after she had an intrusive memory, she reported the intervention to be "very good," and when asked if she found the intervention helpful, she said, "Yes, I forgot time and place and the memory went away immediately."

Self-report Measures on PTSD, Depression and Anxiety Symptoms, and General Functioning

Initial high levels of PTSD symptoms (a PCL-5 score of 51) were reduced by over half at postintervention, and the reduction

was clearly clinically significant at the 3-month follow-up, with a score of only 6 [38]. Depression symptoms were reduced from moderate levels (PHQ-9; 10-14) at baseline to mild (5-9) postintervention, indicating a clinically significant change [40]. Depression symptoms were further reduced to minimal (0-4) at the 3-month follow-up. At baseline, the participant reported mild levels of anxiety (GAD-7; 5-10) and did not report a clinically significant change in symptoms until the 3-month follow-up, when her symptoms were reduced to little or no anxiety (GAD-7; 0-4) [43]. Functional impairment (as measured by the SDS) improved clinically significantly in the follow-up period [45]. The score was 15 at baseline and reduced to zero at the 3-month follow-up (Table 3).

Table 3. Self-report measures for secondary outcomes (posttraumatic stress disorder, depression and anxiety symptoms, and general functioning) and impact of intrusive memories on concentration, sleep, stress, and daily functioning (n=1).

Item	Baseline interview	Postintervention	1-month follow-up	3-month follow-up
PCL-5 ^a	51	35	30	6
PHQ-9 ^b	13	7	8	2
GAD-7 ^c	9	7	7	2
SDS ^d	15	15	5	0
Concentration ^e	5	3	3	1
General concentration ^f	7	3	5	3
Duration of disruption ^g	4	2	2	1
Sleep ^h	5	3	2	0
Nightmares ⁱ	4	6	2	0
Stress ^j	5	3	3	1
Daily functioning ^k	5	1	3	0

^aPCL-5: Posttraumatic Stress Disorder Checklist; scores ranging from 0 to 80.

^bPHQ-9: Patient Health Questionnaire-9; scores ranging from 0 to 27.

^cGAD-7: Generalized Anxiety Disorder scale-7; scores ranging from 0 to 21.

^dSDS: Sheehan Disability Scale; scores ranging from 0 (unimpaired) to 30 (highly impaired).

^eIn the past week, how much did your intrusive memories disrupt your concentration? 0=not at all disruptive; 10=extremely disruptive.

^fIn the past week, how much difficulty did you have concentrating generally? 0=no concentration difficulty at all; 10=extreme concentration difficulty.

^gWhen you had an intrusive memory, how long did it disrupt your concentration (in minutes) in the past week? 0 (<1 minutes) to 5 (>60 minutes).

^hDid your intrusive memories interfere with sleep during the night in the past week? 0=not at all; 10=interfered very much.

ⁱDid you experience any nightmares that interfered with your sleep during the night in the past week? 0=did not experience any nightmares; 10=experienced many nightmares.

^jIn the past week, did your intrusive memories affect how stressed you felt? 0=not at all; 10=affected very much.

^kHave the intrusive memories affected your ability to function in your daily life? 0=not at all; 10=very much affected.

Impact of Intrusive Memories on Concentration, Sleep, Stress, and Daily Functioning

Table 3 shows ratings of the impact of intrusions on concentration, sleep, and stress. Critically, the impact of intrusive memories on *concentration* reduced from 5 at baseline to 1 at the 3-month follow-up, and estimated duration of concentration disruption per intrusion reduced from 4 (30-60 minutes) at baseline to 1 (1-5 minutes) at follow-up. The impact intrusions had on sleep reduced from 5 at baseline to 0 at the 3-month follow-up. The impact intrusions had on stress reduced from 5 at baseline to 1 at the 3-month follow-up. The impact intrusive memories had on the participant's ability to function in her daily life reduced from 5 at baseline to 1 postintervention and was 0 at the 3-month follow-up.

At baseline, the participant responded to an open question on how her intrusive memories had affected her ability to function in daily life: "I don't sleep very well, and that leads to fatigue which interferes with my daily functioning." In the last intervention session, she said, "It took some energy to try not to think about them, but they bother me very little anymore," and at the 1-month follow-up she reported, "I can't concentrate when I have an intrusive memory, but the memories don't really bother me anymore even though I have been in quarantine.

Usually when I am not busy that has meant more memories." She also said, "I have not needed to play Tetris, but it's nice to know that I can if I have an intrusive memory." At the 3-month follow-up, she responded, "They have not been bothering me in the past weeks. It is a little uncomfortable that they may come, but they bother me very little."

Discussion

Principal Findings

In this single case study, we investigated the effects of a brief visuospatial intervention designed to disrupt memory reconsolidation, thereby reducing the number of intrusive memories of trauma. Different intrusive memories were targeted one by one over six sessions, guided by a clinical psychologist. The intervention stemmed from earlier laboratory studies [18,47] as well as clinical studies [19,25]. The total number of intrusive memories per week (primary outcome) was approximately halved from baseline to the intervention phase, similar to what Kessler et al [19] found in a study involving inpatients with complex PTSD. Of particular interest in this study is that the reduction in the number of intrusions continued to 76% at the 1-month follow-up and to 92% at the 3-month follow-up, meaning that three of the four intrusions were eliminated entirely

at 3 months. This critically extends previous studies by examining the *long-term effects* at 3 months postintervention and, in this case, at least suggests that symptoms may continue to improve in the long term rather than rebound. This is perhaps because of the fact that the intervention is simple to use independently once it has been learned so that the participant can self-administer booster doses if needed.

The specific symptom probe design allowed us to zoom in on the effect of each intervention session on each of the participant's four intrusive memories. All four memories reduced after being targeted, with reductions ranging between 26% and 86% (from baseline to intervention phase). By the 3-month follow-up, only the most distressing intrusive memory (*memory A*) was still present, occurring only once during the past week. This quantitative reduction was mirrored in the participant's qualitative feedback, with her noting that the intrusive memories bothered her *very little* at this time.

Symptoms of PTSD (subthreshold for this participant) were reduced postintervention, and the same pattern was observed for symptoms of both depression and anxiety. This change was similar to the results reported by Kessler et al [19]. Interestingly, symptoms of PTSD, depression, and anxiety continued to decrease along with the number of intrusive memories and were minimal at follow-up.

The intrusive memories affected the participant's general functioning at baseline, for example, it affected her sleep, leading to fatigue, which affected her daily functioning. After the intervention, her functioning improved as the intrusions no longer interfered with her day-to-day life at the 3-month follow-up. Her concentration improved considerably from baseline to postintervention and further at follow-up. The participant, in effect, gained back hours during which her concentration was not disrupted by intrusive memories. Both sleep and stress improved postintervention and continued to improve at follow-up.

Importantly, the participant found the gameplay to be a very acceptable way to reduce the frequency of intrusive memories, similar to the Holmes et al [31] study with refugees. The participant also indicated that the intrusive memory diary was straightforward and not burdensome to complete. Most diary

data were successfully recorded, although some days in the baseline and intervention phases were missing, mostly because of extra days passing in between sessions where the participant had not received the next diary provided in sessions.

This intervention approach (currently under development, not evidence based), intended not to treat the whole of PTSD but rather a single symptom, is unlike existing treatment options and potentially removes some common barriers to them. For example, barriers include a sparse number of qualified psychological therapists in Iceland (particularly in rural areas) as well as the high cost of treatment and high dropout rates, stigma, and patients' reluctance to talk about the traumatic experience [10-12].

This intervention removes patients' need to talk about and describe the trauma in detail, is low cost, and because of its simplicity, it may be delivered by nonexperts after training. It is important to explore further how this intervention approach can address other common barriers in existing treatments. Future research should explore remote delivery of the intervention (eg, communication via web-based platforms) instead of in-person meetings [48]. This would remove geographical constraints and make it possible to reach people even when immobilized or isolated (eg, in quarantine because of the COVID-19 pandemic), which is increasingly important in today's uncertain circumstances [49].

Conclusions

Overall, the results of this single case study indicate that the intervention is promising, showing initial signs of effectiveness in reducing the frequency of intrusive memories of trauma that had occurred 4 decades ago and improving mental health and functioning in an Icelandic setting at least for the first participant. The intervention was well tolerated and acceptable, and the effects of the intervention may even continue after the intervention phase. The next step will be to examine whether such effects extend to other participants (eg, in a case series) and to explore remote delivery of the intervention, to explore whether it is possible to deliver by *nonclinicians*, and to further tailor the intervention to this setting based on feedback from target users.

Acknowledgments

The authors would like to thank Johanna Motilla Hoppe and Elin Sjöfn Stephensen for their administrative assistance. EAH received funding from the Oak Foundation (OCAY-18-442) and from the Swedish Research Council (2020-00873) in support of this study; EAH also received funding from AFA Insurance (200342) and the Lupina Foundation. AB received funding from the Icelandic Research Fund (11709-0270). UV received funding to establish the stress and gene analysis (SAGA) cohort from the European Research Council (StressGene, grant 726413) and the Icelandic Research Fund (grant 163362-051). EBT reports funding from the Icelandic Research Fund (185287-051). LS received funding from the Swiss National Science Foundation (P2BEP1_184378) and a Thunberg Fellowship from the Swedish Collegium for Advanced Study.

Conflicts of Interest

EAH reports serving on the board of the charity MQ: Transforming Mental Health (UK). She also receives book royalties from Oxford University Press (Imagery and Cognitive Therapy) and Guilford Press (Imagery-Based Cognitive Therapy for Bipolar Disorder and Mood Instability) and occasional fees from clinical workshops and conference keynotes.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders Internet. Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
2. Kessler RC, Aguilar-Gaxiola S, Alonso J, Benjet C, Bromet EJ, Cardoso G, et al. Trauma and PTSD in the WHO World Mental Health Surveys. *Eur J Psychotraumatol* 2017 Oct 27;8(sup5):1353383 [FREE Full text] [doi: [10.1080/20008198.2017.1353383](https://doi.org/10.1080/20008198.2017.1353383)] [Medline: [29075426](https://pubmed.ncbi.nlm.nih.gov/29075426/)]
3. Post-traumatic stress disorder. National Institute for Health and Care Excellence. 2018 Dec. URL: <https://www.nice.org.uk/guidance/ng116> [accessed 2021-08-31]
4. Santiago PN, Ursano RJ, Gray CL, Pynoos RS, Spiegel D, Lewis-Fernandez R, et al. A systematic review of PTSD prevalence and trajectories in DSM-5 defined trauma exposed populations: intentional and non-intentional traumatic events. *PLoS One* 2013 Apr 11;8(4):e59236 [FREE Full text] [doi: [10.1371/journal.pone.0059236](https://doi.org/10.1371/journal.pone.0059236)] [Medline: [23593134](https://pubmed.ncbi.nlm.nih.gov/23593134/)]
5. Kupfer DJ, Regier DA. Neuroscience, clinical evidence, and the future of psychiatric classification in DSM-5. *Am J Psychiatry* 2011 Jul;168(7):672-674. [doi: [10.1176/appi.ajp.2011.11020219](https://doi.org/10.1176/appi.ajp.2011.11020219)] [Medline: [21724672](https://pubmed.ncbi.nlm.nih.gov/21724672/)]
6. Morina N, Wicherts JM, Lobrecht J, Priebe S. Remission from post-traumatic stress disorder in adults: a systematic review and meta-analysis of long term outcome studies. *Clin Psychol Rev* 2014 Apr;34(3):249-255. [doi: [10.1016/j.cpr.2014.03.002](https://doi.org/10.1016/j.cpr.2014.03.002)] [Medline: [24681171](https://pubmed.ncbi.nlm.nih.gov/24681171/)]
7. McLaughlin KA, Koenen KC, Friedman MJ, Ruscio AM, Karam EG, Shahly V, et al. Subthreshold posttraumatic stress disorder in the world health organization world mental health surveys. *Biol Psychiatry* 2015 Feb 15;77(4):375-384 [FREE Full text] [doi: [10.1016/j.biopsych.2014.03.028](https://doi.org/10.1016/j.biopsych.2014.03.028)] [Medline: [24842116](https://pubmed.ncbi.nlm.nih.gov/24842116/)]
8. Bradley R, Greene J, Russ E, Dutra L, Westen D. A multidimensional meta-analysis of psychotherapy for PTSD. *Am J Psychiatry* 2005 Feb;162(2):214-227. [doi: [10.1176/appi.ajp.162.2.214](https://doi.org/10.1176/appi.ajp.162.2.214)] [Medline: [15677582](https://pubmed.ncbi.nlm.nih.gov/15677582/)]
9. Kantor V, Knepel M, Lueger-Schuster B. Perceived barriers and facilitators of mental health service utilization in adult trauma survivors: a systematic review. *Clin Psychol Rev* 2017 Mar;52:52-68 [FREE Full text] [doi: [10.1016/j.cpr.2016.12.001](https://doi.org/10.1016/j.cpr.2016.12.001)] [Medline: [28013081](https://pubmed.ncbi.nlm.nih.gov/28013081/)]
10. Lewis C, Roberts NP, Gibson S, Bisson JI. Dropout from psychological therapies for post-traumatic stress disorder (PTSD) in adults: systematic review and meta-analysis. *Eur J Psychotraumatol* 2020 Mar 9;11(1):1709709 [FREE Full text] [doi: [10.1080/20008198.2019.1709709](https://doi.org/10.1080/20008198.2019.1709709)] [Medline: [32284816](https://pubmed.ncbi.nlm.nih.gov/32284816/)]
11. Imel ZE, Laska K, Jakupcak M, Simpson TL. Meta-analysis of dropout in treatments for posttraumatic stress disorder. *J Consult Clin Psychol* 2013 Jun;81(3):394-404 [FREE Full text] [doi: [10.1037/a0031474](https://doi.org/10.1037/a0031474)] [Medline: [23339535](https://pubmed.ncbi.nlm.nih.gov/23339535/)]
12. Najavits LM. The problem of dropout from "gold standard" PTSD therapies. *F1000Prime Rep* 2015 Apr 2;7:43 [FREE Full text] [doi: [10.12703/P7-43](https://doi.org/10.12703/P7-43)] [Medline: [26097716](https://pubmed.ncbi.nlm.nih.gov/26097716/)]
13. Kessler RC. Posttraumatic stress disorder: the burden to the individual and to society. *J Clin Psychiatry* 2000;61 Suppl 5:4-12; discussion 13. [Medline: [10761674](https://pubmed.ncbi.nlm.nih.gov/10761674/)]
14. Singh L, Espinosa L, Ji JL, Moulds ML, Holmes EA. Developing thinking around mental health science: the example of intrusive, emotional mental imagery after psychological trauma. *Cogn Neuropsychiatry* 2020 Sep;25(5):348-363. [doi: [10.1080/13546805.2020.1804845](https://doi.org/10.1080/13546805.2020.1804845)] [Medline: [32847486](https://pubmed.ncbi.nlm.nih.gov/32847486/)]
15. Ehlers A, Hackmann A, Michael T. Intrusive re-experiencing in post-traumatic stress disorder: phenomenology, theory, and therapy. *Memory* 2004 Jul;12(4):403-415. [doi: [10.1080/09658210444000025](https://doi.org/10.1080/09658210444000025)] [Medline: [15487537](https://pubmed.ncbi.nlm.nih.gov/15487537/)]
16. Grey N, Holmes EA. "Hotspots" in trauma memories in the treatment of post-traumatic stress disorder: a replication. *Memory* 2008 Oct;16(7):788-796. [doi: [10.1080/09658210802266446](https://doi.org/10.1080/09658210802266446)] [Medline: [18720224](https://pubmed.ncbi.nlm.nih.gov/18720224/)]
17. Iyadurai L, Visser RM, Lau-Zhu A, Porcheret K, Horsch A, Holmes EA, et al. Intrusive memories of trauma: a target for research bridging cognitive science and its clinical application. *Clin Psychol Rev* 2019 Apr;69:67-82 [FREE Full text] [doi: [10.1016/j.cpr.2018.08.005](https://doi.org/10.1016/j.cpr.2018.08.005)] [Medline: [30293686](https://pubmed.ncbi.nlm.nih.gov/30293686/)]
18. Holmes EA, James EL, Coode-Bate T, Deerprouse C. Can playing the computer game "Tetris" reduce the build-up of flashbacks for trauma? A proposal from cognitive science. *PLoS One* 2009;4(1):e4153 [FREE Full text] [doi: [10.1371/journal.pone.0004153](https://doi.org/10.1371/journal.pone.0004153)] [Medline: [19127289](https://pubmed.ncbi.nlm.nih.gov/19127289/)]
19. Kessler H, Holmes EA, Blackwell SE, Schmidt A, Schweer JM, Bücker A, et al. Reducing intrusive memories of trauma using a visuospatial interference intervention with inpatients with posttraumatic stress disorder (PTSD). *J Consult Clin Psychol* 2018 Dec;86(12):1076-1090. [doi: [10.1037/ccp0000340](https://doi.org/10.1037/ccp0000340)] [Medline: [30507232](https://pubmed.ncbi.nlm.nih.gov/30507232/)]
20. Hoge CW, Chard KM. A window into the evolution of trauma-focused psychotherapies for posttraumatic stress disorder. *JAMA* 2018 Jan 23;319(4):343-345. [doi: [10.1001/jama.2017.21880](https://doi.org/10.1001/jama.2017.21880)] [Medline: [29362777](https://pubmed.ncbi.nlm.nih.gov/29362777/)]
21. Holmes E, Hales S, Young K, Simplicio M. *Imagery-Based Cognitive Therapy for Bipolar Disorder and Mood Instability*. New York: Guilford Press; 2019.
22. Iyadurai L, Blackwell SE, Meiser-Stedman R, Watson PC, Bonsall MB, Geddes JR, et al. Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: a proof-of-concept randomized controlled trial. *Mol Psychiatry* 2018 Mar;23(3):674-682 [FREE Full text] [doi: [10.1038/mp.2017.23](https://doi.org/10.1038/mp.2017.23)] [Medline: [28348380](https://pubmed.ncbi.nlm.nih.gov/28348380/)]

23. Horsch A, Vial Y, Favrod C, Harari MM, Blackwell SE, Watson P, et al. Reducing intrusive traumatic memories after emergency caesarean section: a proof-of-principle randomized controlled study. *Behav Res Ther* 2017 Jul;94:36-47 [FREE Full text] [doi: [10.1016/j.brat.2017.03.018](https://doi.org/10.1016/j.brat.2017.03.018)] [Medline: [28453969](https://pubmed.ncbi.nlm.nih.gov/28453969/)]
24. Kanstrup M, Singh L, Göransson KE, Widoff J, Taylor RS, Gamble B, et al. Reducing intrusive memories after trauma via a brief cognitive task intervention in the hospital emergency department: an exploratory pilot randomised controlled trial. *Transl Psychiatry* 2021 Jan 11;11(1):30 [FREE Full text] [doi: [10.1038/s41398-020-01124-6](https://doi.org/10.1038/s41398-020-01124-6)] [Medline: [33431807](https://pubmed.ncbi.nlm.nih.gov/33431807/)]
25. Kanstrup M, Kontio E, Geranmayeh A, Olofsdotter Lauri K, Moulds ML, Holmes EA. A single case series using visuospatial task interference to reduce the number of visual intrusive memories of trauma with refugees. *Clin Psychol Psychother* 2021 Jan;28(1):109-123. [doi: [10.1002/cpp.2489](https://doi.org/10.1002/cpp.2489)] [Medline: [32525244](https://pubmed.ncbi.nlm.nih.gov/32525244/)]
26. Iyadurai L, Hales SA, Blackwell SE, Young K, Holmes EA. Targeting intrusive imagery using a competing task technique: a case study. *Behav Cogn Psychother* 2020 Nov;48(6):739-744. [doi: [10.1017/S1352465820000296](https://doi.org/10.1017/S1352465820000296)] [Medline: [32594968](https://pubmed.ncbi.nlm.nih.gov/32594968/)]
27. Krasny-Pacini A, Evans J. Single-case experimental designs to assess intervention effectiveness in rehabilitation: a practical guide. *Ann Phys Rehabil Med* 2018 May;61(3):164-179 [FREE Full text] [doi: [10.1016/j.rehab.2017.12.002](https://doi.org/10.1016/j.rehab.2017.12.002)] [Medline: [29253607](https://pubmed.ncbi.nlm.nih.gov/29253607/)]
28. Vohra S, Punja S. A case for n-of-1 trials. *JAMA Intern Med* 2019 Mar 01;179(3):452. [doi: [10.1001/jamainternmed.2018.7166](https://doi.org/10.1001/jamainternmed.2018.7166)] [Medline: [30830183](https://pubmed.ncbi.nlm.nih.gov/30830183/)]
29. Senn S. Statistical pitfalls of personalized medicine. *Nature* 2018 Nov;563(7733):619-621. [doi: [10.1038/d41586-018-07535-2](https://doi.org/10.1038/d41586-018-07535-2)] [Medline: [30482931](https://pubmed.ncbi.nlm.nih.gov/30482931/)]
30. Bradbury J, Avila C, Grace S. Practice-based research in complementary medicine: could N-of-1 trials become the new gold standard? *Healthcare (Basel)* 2020 Jan 08;8(1):15 [FREE Full text] [doi: [10.3390/healthcare8010015](https://doi.org/10.3390/healthcare8010015)] [Medline: [31936355](https://pubmed.ncbi.nlm.nih.gov/31936355/)]
31. Holmes EA, Ghaderi A, Eriksson E, Lauri KO, Kukacka OM, Mamish M, et al. 'I Can't Concentrate': a feasibility study with young refugees in Sweden on developing science-driven interventions for intrusive memories related to trauma. *Behav Cogn Psychother* 2017 Mar;45(2):97-109 [FREE Full text] [doi: [10.1017/S135246581600062X](https://doi.org/10.1017/S135246581600062X)] [Medline: [28229806](https://pubmed.ncbi.nlm.nih.gov/28229806/)]
32. Tetris holding internet. Electronic Arts Inc. 2018. URL: <https://www.ea.com/de-de/games/tetris/tetris> [accessed 2021-08-31]
33. 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/)]
34. Weathers FW, Bovin MJ, Lee DJ, Sloan DM, Schnurr PP, Kaloupek DG, et al. The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans. *Psychol Assess* 2018 Mar;30(3):383-395 [FREE Full text] [doi: [10.1037/pas0000486](https://doi.org/10.1037/pas0000486)] [Medline: [28493729](https://pubmed.ncbi.nlm.nih.gov/28493729/)]
35. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan K, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur psychiatr* 1997;12(5):224-231. [doi: [10.1016/S0924-9338\(97\)83296-8](https://doi.org/10.1016/S0924-9338(97)83296-8)]
36. Sheehan D, Lecrubier Y, Harnett Sheehan K, Janavs J, Weiller E, Keskiner A, et al. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *Eur psychiatr* 1997;12(5):232-241. [doi: [10.1016/s0924-9338\(97\)83297-x](https://doi.org/10.1016/s0924-9338(97)83297-x)]
37. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. *J Trauma Stress* 2015 Dec;28(6):489-498. [doi: [10.1002/jts.22059](https://doi.org/10.1002/jts.22059)] [Medline: [26606250](https://pubmed.ncbi.nlm.nih.gov/26606250/)]
38. Wortmann JH, Jordan AH, Weathers FW, Resick PA, Dondanville KA, Hall-Clark B, et al. Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychol Assess* 2016 Nov;28(11):1392-1403. [doi: [10.1037/pas0000260](https://doi.org/10.1037/pas0000260)] [Medline: [26751087](https://pubmed.ncbi.nlm.nih.gov/26751087/)]
39. 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/)]
40. Kroenke K. Enhancing the clinical utility of depression screening. *CMAJ* 2012 Feb 21;184(3):281-282 [FREE Full text] [doi: [10.1503/cmaj.112004](https://doi.org/10.1503/cmaj.112004)] [Medline: [22231681](https://pubmed.ncbi.nlm.nih.gov/22231681/)]
41. 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/)]
42. 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/)]
43. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). *J Affect Disord* 2020 Mar 15;265:395-401. [doi: [10.1016/j.jad.2020.01.032](https://doi.org/10.1016/j.jad.2020.01.032)] [Medline: [32090765](https://pubmed.ncbi.nlm.nih.gov/32090765/)]
44. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *Int J Psychiatry Med* 1997;27(2):93-105. [doi: [10.2190/T8EM-C8YH-373N-UWU](https://doi.org/10.2190/T8EM-C8YH-373N-UWU)] [Medline: [9565717](https://pubmed.ncbi.nlm.nih.gov/9565717/)]

45. Coles T, Coon C, DeMuro C, McLeod L, Gnanasakthy A. Psychometric evaluation of the Sheehan Disability Scale in adult patients with attention-deficit/hyperactivity disorder. *Neuropsychiatr Dis Treat* 2014 May 19;10:887-895 [FREE Full text] [doi: [10.2147/NDT.S55220](https://doi.org/10.2147/NDT.S55220)] [Medline: [24899807](https://pubmed.ncbi.nlm.nih.gov/24899807/)]
46. Bjornsson AS, Hardarson JP, Valdimarsdottir AG, Gudmundsdottir K, Tryggvadottir A, Thorarinsdottir K, et al. Social trauma and its association with posttraumatic stress disorder and social anxiety disorder. *J Anxiety Disord* 2020 May;72:102228 [FREE Full text] [doi: [10.1016/j.janxdis.2020.102228](https://doi.org/10.1016/j.janxdis.2020.102228)] [Medline: [32361167](https://pubmed.ncbi.nlm.nih.gov/32361167/)]
47. James EL, Bonsall MB, Hoppitt L, Tunbridge EM, Geddes JR, Milton AL, et al. Computer game play reduces intrusive memories of experimental trauma via reconsolidation-update mechanisms. *Psychol Sci* 2015 Aug;26(8):1201-1215 [FREE Full text] [doi: [10.1177/0956797615583071](https://doi.org/10.1177/0956797615583071)] [Medline: [26133572](https://pubmed.ncbi.nlm.nih.gov/26133572/)]
48. Gamble B, Depa K, Holmes EA, Kanstrup M. Digitalizing a brief intervention to reduce intrusive memories of psychological trauma: qualitative interview study. *JMIR Ment Health* 2021 Feb 22;8(2):e23712 [FREE Full text] [doi: [10.2196/23712](https://doi.org/10.2196/23712)] [Medline: [33616540](https://pubmed.ncbi.nlm.nih.gov/33616540/)]
49. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry* 2020 Jun;7(6):547-560 [FREE Full text] [doi: [10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1)] [Medline: [32304649](https://pubmed.ncbi.nlm.nih.gov/32304649/)]

Abbreviations

CAPS: Clinician Administered PTSD Scale
CONSORT: Consolidated Standards of Reporting Trials
DSM: *Diagnostic and Statistical Manual of Mental Disorders*
GAD-7: Generalized Anxiety Disorder-7
MINI: Mini International Neuropsychiatric Interview
PCL-5: Posttraumatic Stress Disorder Checklist-5
PHQ-9: Patient Health Questionnaire-9
PTSD: posttraumatic stress disorder
REDCap: Research Electronic Data Capture
SAGA: stress and gene analysis
SDS: Sheehan Disability Scale

Edited by G Eysenbach; submitted 23.04.21; peer-reviewed by J Hoogervorst; comments to author 20.05.21; revised version received 11.06.21; accepted 05.07.21; published 04.11.21.

Please cite as:

Thorarinsdottir K, Holmes EA, Hardarson J, Hedinsdottir U, Kanstrup M, Singh L, Hauksdottir A, Halldorsdottir T, Gudmundsdottir B, Valdimarsdottir U, Thordardottir EB, Gamble B, Bjornsson A
Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland
JMIR Form Res 2021;5(11):e29873
URL: <https://formative.jmir.org/2021/11/e29873>
doi: [10.2196/29873](https://doi.org/10.2196/29873)
PMID: [34734830](https://pubmed.ncbi.nlm.nih.gov/34734830/)

©Kristjana Thorarinsdottir, Emily A Holmes, Johann Hardarson, Unnur Hedinsdottir, Marie Kanstrup, Laura Singh, Arna Hauksdottir, Thorhildur Halldorsdottir, Berglind Gudmundsdottir, Unnur Valdimarsdottir, Edda Bjork Thordardottir, Beau Gamble, Andri Bjornsson. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 04.11.2021. 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 Email, SMS, and Concurrent Mixed Modes Approaches to Capture Quality of Recovery in the Perioperative Period: Retrospective Longitudinal Cohort Study

Jamie L Romeiser¹, MPH, PhD; James Cavalcante¹, BS, MD; Deborah C Richman¹, MBChB, FFA(SA); Sunitha M Singh¹, MD; Xiaohui Liang¹, BS; Allison Pei¹, BS; Samanvaya Sharma¹, BA; Zoe Lazarus¹, BS; Tong J Gan¹, MD; Elliott Bennett-Guerrero¹, MD

Department of Anesthesiology, Stony Brook University Medical Center, Stony Brook, NY, United States

Corresponding Author:

Jamie L Romeiser, MPH, PhD

Department of Anesthesiology

Stony Brook University Medical Center

101 Nicolls Rd

Stony Brook, NY, 11794

United States

Phone: 1 631 444 2975

Email: jamie.romeiser@stonybrookmedicine.edu

Abstract

Background: As patients are discharged from the hospital more quickly, the ability to monitor patient recovery between hospital discharge and the first follow-up clinic visit is becoming increasingly important. Despite substantial increase in both internet use and smartphone ownership over the past 5 years, clinicians have been slow to embrace the use of these devices to capture patient recovery information in the period between hospital discharge and the first clinical follow-up appointment.

Objective: This study aims to investigate the generalizability of using a web-based platform to capture patient recovery in a broad surgical patient population and compare response rates for 3 different web-based strategies for delivering recovery surveys over the perioperative period: email, SMS text messaging, and a concurrent mixed approach of using both email and SMS text messaging.

Methods: Patients undergoing surgeries managed with an enhanced recovery after surgery pathway were asked to participate in a web-based quality assurance monitoring program at the time of their preoperative surgery appointment. Different follow-up methods were implemented over 3 sequential phases. Patients received Health Insurance Portability and Accountability Act-compliant web-based survey links via email (phase 1), SMS text messaging (phase 2), or concurrently using both email and SMS text messaging (phase 3) using REDCap and Twilio software. Recovery assessments using the established Quality of Recovery-9 instrument were performed 4 days before surgery and at 7 and 30 days postoperatively. Generalizability of the web-based system was examined by comparing characteristics of those who participated versus those who did not. Differences in response rates by the web-based collection method were analyzed using adjusted models.

Results: A total of 615 patients were asked to participate, with 526 (85.5%) opting for the follow-up program. Those who opted in were younger, slightly healthier, and more likely to be in a partnership. The concurrent mixed modes method was the most successful for obtaining responses at each time point compared with text or email alone (pre: 119/160, 74.4% vs 116/173, 67.1% vs 56/130, 43.1%, $P<.001$; 7 days: 115/172, 66.9% vs 82/164, 50.0% vs 59/126, 46.8%, $P=.001$; 30 days: 152/234, 65.0% vs 52/105, 49.5% vs 53/123, 43.1%, $P=.001$, respectively). In the adjusted model, the concurrent mixed modes method significantly predicted response compared with using email alone (odds ratio 3.4; $P<.001$) and SMS text messaging alone (odds ratio 1.9; $P<.001$). Additional significant predictors of response were race, partnership, and time.

Conclusions: For internet users and smartphone owners, electronic capture of recovery surveys appear to be possible through this mechanism. Discrepancies in both inclusion and response rates still exist among certain subgroups of patients, but the concurrent approach of using both email and text messages was the most effective approach to reach the largest number of patients across all subgroups.

(JMIR Form Res 2021;5(11):e25209) doi:[10.2196/25209](https://doi.org/10.2196/25209)

KEYWORDS

concurrent mixed modes; recovery after surgery; text messages; SMS; email; perioperative recovery; mobile phone

Introduction

Background

Patients are discharged from the hospital more quickly with the help of minimally invasive procedures, safer and shorter acting drugs, enhanced recovery programs, and advances in patient safety and monitoring [1,2]. There are several benefits of a reduced length of hospital stay, including reductions in nosocomial infections [3], opioid use, and medical expenditures [2]. However, this also means that a larger proportion of the patient's recovery process is spent at home. What happens in the short-term recovery process, that is, the time between the time of discharge and the first clinical follow-up appointment, can be crucial for setting the trajectory of a patient's long-term recovery after surgery [4].

Patient-reported outcomes (PRO) are an important component of clinical outcomes [5-7]. The collection of PROs, especially once a patient has been discharged, is challenging and is thus underutilized in clinical practice [8]. Finding a practical and affordable way to contact patients outside of a clinical setting is challenging. Sending postal surveys leads to a large delay in the return of information, and repeated telephone calls require a significant amount of time and resources. A more practical solution would be to use a web-based approach. It is estimated that 90% of adults in the United States use the internet on a regular basis, and 81% own a smartphone [9].

The use of this technology is at an all-time high in the United States, but it is not often leveraged to implement short-term postdischarge recovery monitoring programs. Patient demographics, such as age [10], can be a barrier for inclusion in recovery monitoring programs that use SMS-text. High response rates have been demonstrated for these types of SMS programs [11-13], but many of these studies did not report the number of patients or demographics of patients who were excluded from the intervention. Other studies that focus on long-term recovery surveys (ie, 3 months-2 years) have found some success with using email methods to deliver these surveys [14,15]. However, these studies used secondary methods, such as telephone reminders and postal questionnaires, to obtain responses. Patients cited lack of email, lack of internet access, infrequent email use, or browser incompatibility as reasons for not completing web-based questionnaires [14], and response rates have been shown to differ by race, household income, and procedural characteristics [15].

Objectives

The ability to maximize participation and engagement to produce generalizable and unbiased information for quality-of-care improvement is essential. One approach that is rarely considered is to send PRO recovery surveys through a concurrent mixed-mode approach, which uses a combination of simultaneous email and text messages. Thorough examination of the generalizability and feasibility of such a system in the short-term perioperative period has not been explored. To our

knowledge, response rates using email alone, text alone, and this combined approach have not been directly compared in this setting. Therefore, the primary aim of this study is to evaluate and compare the feasibility of collecting short-term perioperative PROs within an enhanced recovery after surgery (ERAS) population using 3 different web-based and mobile phone collection mechanisms—email alone, text alone, and a concurrent mixed modes (CMM) approach.

Methods

Study Design and Participants

This project follows the Standards for Quality Improvement Reporting Excellence (SQUIRE) guidelines. Although this project was authorized by the institution's Division of Medical and Regulatory Affairs Office as part of its Surgical Quality Improvement Program, following deidentification and extraction of all Surgical Quality Improvement Program patient records, data analyses were performed as part of the Surgical Quality Data Users Group. The Surgical Quality Improvement Program/Surgical Quality Data Users Group protocols were approved by our institution's review board (Committee on Research Involving Human Subjects, CORIHS #170753-13).

The longitudinal cohort consisted of surgical patients managed by an ERAS pathway. This quality assurance (QA) program includes patients who are undergoing colorectal, minimally invasive gynecology, orthopedic, plastic, lumbar spine, urology, surgical oncology, and urogynecology-related surgeries. In 2018, the program implemented a web-based element in which patients could opt-in and complete follow-up questionnaires regarding their recovery after surgery. All patients who were treated using the ERAS pathway were eligible to participate in the follow-up portion of the QA program. Opt-in forms were completed by the patients during their preoperative services visit.

Survey Instrument, Software, and Distribution

We used the Quality of Recovery-9 (QOR-9) survey, which was developed specifically to quantify overall recovery status after surgery and anesthesia [16,17], and has been validated in a diverse set of surgical patients both preoperatively and postoperatively at time points ranging from 1 day to 6 weeks [17]. REDCap (Vanderbilt University) software with Twilio (Twilio Inc) application integration was used for the survey distribution. Twilio is a separate third-party app that can send text messages with survey links. We used the social exchange theory to guide several survey design choices, such as including a welcome message with information about patient privacy, adding an institutional logo, using a QA-specific email address, and ensuring that the survey was brief. Surveys were compatible with all types of browsers and devices, and fonts and colors were accessible to individuals who may have had a vision impairment.

Automated surveys were programmed to be sent to patients based on their procedure date. Surveys were sent through Health

Insurance Portability and Accountability Act–compliant survey links, even though surveys did not contain any protected health information or personally identifiable information. Three time points were chosen to administer the QOR-9 survey: 4 days before the patient's planned procedure date and at 7 and 30 days after the planned procedure date. Surveys were delivered by a different method of contact across 3 sequential phases over a 6-month time span. Patients who underwent surgery in the first phase were preassigned to receive the surveys by email link alone; patients who had surgery in the second phase were preassigned to receive the surveys by a text link alone; and patients who had surgery in the third phase were preassigned to receive the surveys through a CMM approach, and they received both an email link and a text link at the same time. If participants only listed one mode of contact (ie, only email or only smartphone number), the survey was sent to that method of contact regardless of the preassignment. This was done to prevent any exclusion from the QA program. Analyses were run as *per protocol* analyses, with the actual method of contact as the primary predictor in the response models.

All surveys were sent at 9:15 AM local time. One automated reminder per survey was programmed into the REDCap software and sent to the patients who had not responded to the original survey within 24 hours. At the preoperative and 7-day postoperative time points, a follow-up phone call was also made if patients did not respond within 24 hours to the automated reminder. All survey links expired after 3 days.

Study Variables

Method of Contact

The primary predictor of response was the method of contact. The method of contact used to send survey links to participants was either through email alone, SMS text messaging alone, or through both an email and SMS simultaneously (CMM).

Outcomes

The primary outcome was response to the survey questionnaires. Because the primary purpose of this study was to examine the feasibility of a web-based approach to obtain recovery information, responses were only considered positive if they were obtained by the web-based system. Responses obtained through telephone calls were characterized as a nonresponse. Response variables were created separately for each time point (preoperatively, 7, and 30 days). Generalizability (ie, inclusion) was a secondary outcome and was defined as opting into the follow-up program versus opting out.

Demographic Characteristics

Individual-level sociodemographic characteristics, including age, race, ethnicity, and partnership (marital status), were extracted from the hospital's electronic medical record for all ERAS QA program participants. Age was examined continuously but was also dichotomized into <75 and ≥75 years old. This dichotomy was chosen as previous literature describes a noted decline in the use of web technology in individuals 75 years of age and above [18]. Partnership was defined as either being currently married or in a partnership. The primary language was categorized as English and non-English speaking.

Clinical Characteristics

Each element of the Charlson comorbidity index (CCI) was collected from the electronic medical record according to the definitions outlined in the index instructions [19]. Age was not included in the index calculation to avoid multicollinearity. A CCI score higher than 3 was rare; therefore, for ease of interpretation and model stability, the index was categorized into 4 groups: 0, 1, 2, and 3+. Additional key comorbidities, such as depression and anxiety (defined as an active diagnosis or depression medication within 3 months of surgery) and chronic pain (defined as an active diagnosis or opioid prescription within 3 months of surgery), were extracted from the preoperative service visit records in the electronic medical record.

Population-Level Characteristics

Zip code level factors, including inflation-adjusted median income and proportion of residents with greater than a high school education, were obtained from the United States Census Bureau's 2017 American Community Survey 5-year estimates.

Statistical Analysis

Demographics and clinical comorbidities were compared between those who opted in and those who opted out of the follow-up program using both univariate and multivariable logistic regression models. Similar analyses were performed between those who responded and those who did not at each time point, with the method of contact as the primary predictor of interest. All time point data were then combined for a final multivariable model using generalized estimating equations with an independent correlation structure to adjust for repeated measures. Multicollinearity was present for education and median income zip code level factors in all analyses; therefore, only median income was used in the multiple regression models. Sensitivity analyses were performed to address any effects that coverage error may have on the nonresponse analyses by using inverse probability weighting. Inverse probability weighting is a technique used to make the response sample (ie, those who opted in) more reflective of the original ERAS population. Weights were calculated on the basis of the probability of inclusion from the adjusted logistic regression model in the generalizability analysis and applied to the *opt-in* population for a weighted generalized estimating equations sensitivity analysis [20-22].

A common concern in survey research is nonresponse bias, as it relates to missing outcomes. One strategy used to overcome this issue is to develop a proxy group for nonresponders, which can be based on the continuum of resistance theory [23,24]. This theory states that late responders have similar characteristics and outcomes as nonresponders [25]. Therefore, we created a variable that indicated *early* responders (ie, response on the first request), *delayed* responders (ie, response to the 24-hour reminder), and *late* responders (ie, those who responded to the telephone call) to examine nonresponse bias. QOR-9 scores were first tested for normality. Data were not normally distributed (Shapiro–Wilk test $P < .05$); therefore, QOR-9 scores between the 3 response groups were compared

using a Kruskal–Wallis test. Tests were performed separately for the preoperative and 7-day postoperative time points.

Three separate Blinder-Oaxaca decomposition analyses were performed to further characterize any differences in the composition of the *email alone* group compared with the *SMS alone* group, CMM compared with the SMS alone group, and CMM compared with the email alone group. Blinder-Oaxaca decomposition analyses were performed using STATA's (version 15, Stata Corp) *mvdcmp* command with logistic regression. All other analyses were performed using SAS version 9.4 (SAS Institute Inc) at 95% CI.

Results

Description of Patient Characteristics

A flow diagram of the participants is shown in [Figure 1](#). A total of 615 patients were eligible: 526 (85.5%) opted-in to the additional follow-up program and 89 (14.5%) either declined or did not have an email address or smartphone. An additional 63 participants were excluded from the response analyses for various reasons. The characteristics of the patients with ERAS are described in [Table 1](#).

Figure 1. Flow diagram of enhanced recovery after surgery participants. Patients who opted in (n=526) and those who opted out (n=89) were compared in the generalizability analyses. Those who opted in and were not excluded for other reasons were examined in the response analyses (n=463). Those who responded after the phone call were not considered responders in the primary response analysis because they did not use the web-based tool to respond; instead, these individuals were included as late responders in the Quality of Recovery score analyses that examined nonresponse bias within outcomes.

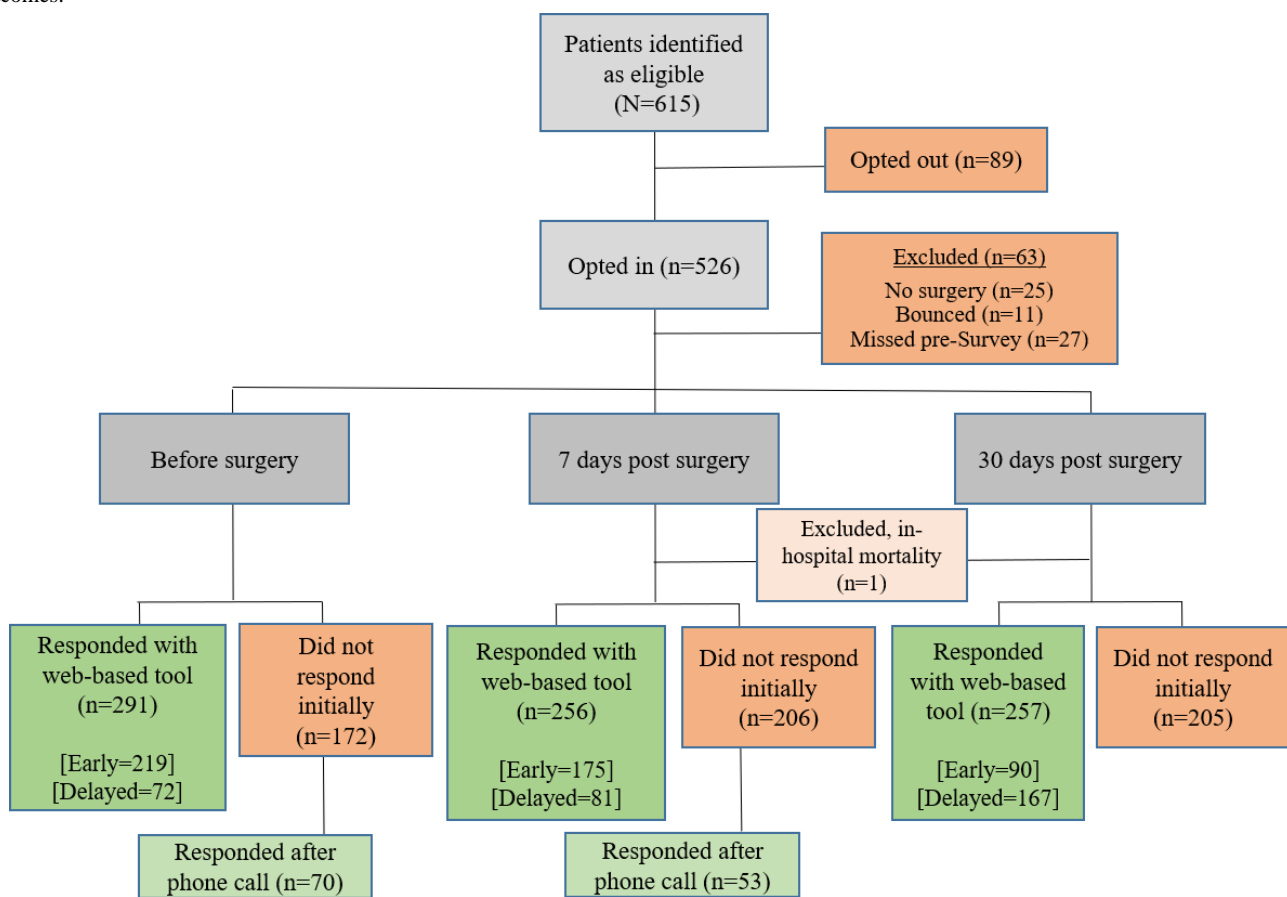


Table 1. Demographic and clinical characteristics of the enhanced recovery after surgery population.

Characteristics	All (N=615)	Opted-in (n=526, 85.5)	Opted-out (n=89, 14.5)	P value
Demographics				
Sex (female), n (%)	404 (65.7)	345 (65.6)	59 (66.3)	.90 ^a
Age (years), mean (SD)	62.6 (12.0)	61.64 (11.5)	68.21 (13.5)	<.001 ^b
Age (≥75 years), n (%)	95 (15.5)	61 (11.6)	34 (38.2)	<.001 ^a
Race/ethnicity (non-Hispanic White), n (%)	533 (86.7)	457 (86.9)	76 (85.4)	.89 ^a
Partnership, n (%)	425 (69.1)	379 (72.1)	46 (51.7)	<.001 ^a
Language (English), n (%)	602 (97.9)	517 (98.3)	85 (95.5)	.10 ^c
Population-level characteristics				
Median income, mean (SD)	\$96,908 (\$20,081)	\$96,889 (\$20,290)	\$97,020 (\$18,868)	.96 ^b
Education (proportion>High School), mean (SD)	0.6 (0.1)	0.6 (0.1)	0.6 (0.1)	.58 ^b
Comorbidities				
Depression or anxiety, n (%)	109 (17.7)	93 (17.7)	16 (18.0)	.95 ^a
Chronic pain, n (%)	39 (6.3)	37 (7.0)	2 (2.3)	.10 ^c
Charlson comorbidity index categories, n (%)				<.001 ^a
0	285 (46.3)	256 (48.7)	29 (32.6)	
1	159 (25.9)	140 (26.6)	19 (21.4)	
2	101 (16.4)	78 (14.8)	23 (25.8)	
>3	70 (11.4)	52 (9.9)	18 (20.2)	

^aAnalyzed using chi-square test.

^bAnalyzed using student *t* test.

^cAnalyzed using Fisher exact test.

Generalizability

In the adjusted analyses, those ≥75 years old were much less likely to participate compared with those aged <75 years (odds ratio [OR] 0.2, 95% CI 0.1-0.4; *P*<.001; [Figure 2](#), [Table 2](#)).

Those with a higher CCI of either 2 or >3 had lower odds of participating compared with those with a CCI score of 0, and those in a current partnership or marriage had more than twice the odds of participating compared with those who were not (OR 2.2, 95% CI 1.3-3.6; *P*=.002).

Figure 2. Forest plot display of the adjusted odds ratios and 95% CIs for characteristics of opting in to the follow-up program. Results from the multivariable logistic regression for characteristics of inclusion were displayed (odds ratio and 95% CI) as a forest plot.

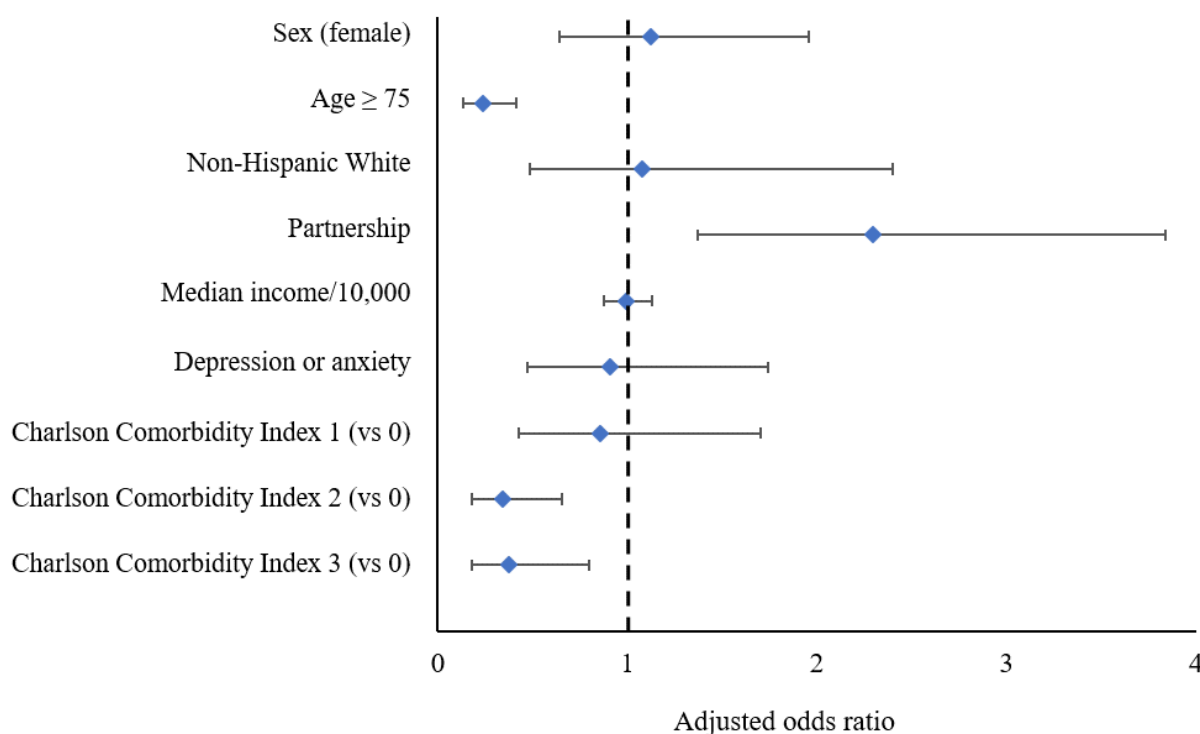


Table 2. Adjusted odds of opting into the follow-up program.

Characteristics of response	OR ^a (95% CI)	P value
Sex (female)	1.1 (0.7-1.9)	.71
Age (≥75 years)	0.2 (0.14-0.4)	<.001
Race/ethnicity (non-Hispanic white)	1.2 (0.6-2.6)	.61
Partnership	2.2 (1.34-3.6)	.002
Median income (per US \$10,000)	1.0 (0.9-1.1)	.92
Depression or anxiety	1.0 (0.5-1.8)	.92
Charlson comorbidity categories		
0	Reference	Reference
1	0.9 (0.5-1.8)	.76
2	0.4 (0.2-0.7)	.003
>3	0.4 (0.2-0.8)	.008

^aOR: odds ratio.

Characteristics of Response

Of the 463 patients included in the response analyses, 291 (62.9%) responded to the preoperative survey, 256 (55.4%) responded to the 7-day postoperative survey, and 257 (55.6%) responded to the 30-day postoperative survey. Notably, response rates significantly differed by method of contact at each time point, with the CMM approach having the highest response rates over time compared with the single-mode approaches (Figure 3). This pattern of response by method of contact was

mostly consistent within several exploratory subgroup analyses (eg, within subgroups of partnership and race), although differences did not reach statistical significance (results not shown). Rates of response were similar by method of contact for the older age group at the pre and 7-day time points, but CMM was superior at the 30-day time point (results not shown).

Multivariable logistic regression results for possible response characteristics are presented for the 3 time points in Table 3, and for all models in Figure 4. In the adjusted models, the

method of contact, race, and partnership were the most consistent determinants of response at each time point. After adjusting for all characteristics in the generalized estimating equations analysis (Figure 4; Table 4), using CMM or SMS alone resulted in significantly higher odds of response compared with the email group (CMM OR 3.4, $P < .001$; SMS OR 1.8, $P = .003$). In contrast, CMM also demonstrated significantly higher adjusted odds of response compared with SMS alone (CMM OR 1.9, $P < .001$). The additional significant predictors were race, partnership, and time.

Inverse probability weights were calculated from the generalizability-adjusted logistic regression model and applied to the nonresponse models in a sensitivity analysis (Figure 4). Most of the effects and significance levels for the characteristics in each time period model remained the same, indicating that the coverage error had little effect on the response analyses. Decomposition analyses revealed no significant differences in the composition (ie, demographic or clinical characteristics) of the CMM, email alone, and SMS alone groups. All differences in the effects on response were attributed solely to the method of contact.

Figure 3. Response rates by method of contact. There were significant differences in response rates according to the method of contact at all 3 time points. The concurrent mixed-mode approach resulted in significantly higher response rates over time compared with single-mode approaches alone. *Chi-square P value for the 3-group comparison $< .001$. **Chi-square P value for the 3-group comparison $= .001$.

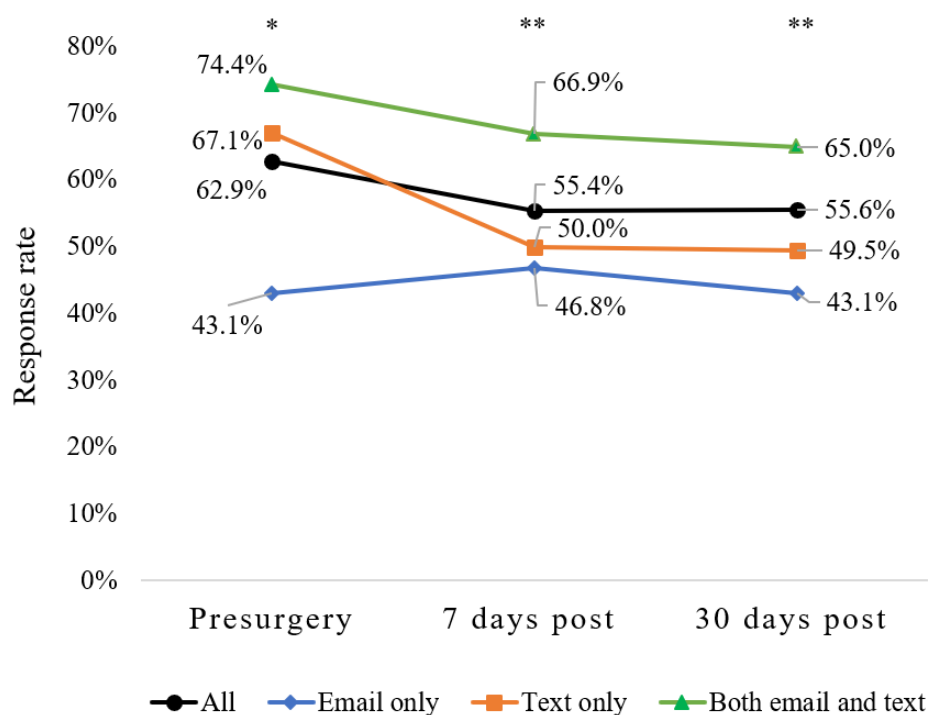


Table 3. Multivariable logistic regression model results for each time point.

Characteristics	Preoperative		7-day postoperative		30-day postoperative	
	OR ^a (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value
Method of contact						
Email	Reference	N/A ^b	Reference	N/A	Reference	N/A
Text	2.9 (2.7-8.1)	<.001	1.3 (0.8-2.2)	.29	1.5 (0.9-2.7)	.14
Both	4.6 (1.7-4.8)	<.001	3.0 (1.8-5.1)	<.001	3.4 (2.0-5.5)	<.001
Sex (female)	0.6 (0.4-1.0)	.06	0.9 (0.6-1.5)	.90	1.1 (0.7-1.8)	.60
Age (≥75 years)	0.7 (0.4-1.4)	.35	0.7 (0.4-1.2)	.18	0.7 (0.4-1.3)	.29
Race ethnicity (White)	5.3 (2.8-10.3)	<.001	2.6 (1.4-4.9)	.003	3.4 (1.8-6.6)	<.001
Partnership	1.2 (0.7-1.8)	.56	1.9 (1.2-2.9)	.006	2.1 (1.3-3.3)	.002
Median income ^c	1.0 (0.9-1.1)	.38	1.1 (0.9-1.2)	.31	1.0 (0.9-1.1)	.62
Depression or anxiety	1.3 (0.7-2.3)	.39	0.7 (0.4-1.1)	.11	0.9 (0.6-1.6)	.88
Chronic pain	0.8 (0.4-2.0)	.70	0.9 (0.4-2.1)	.81	0.3 (0.1-0.7)	.01
Charlson comorbidity index						
0	Reference	N/A	Reference	N/A	Reference	N/A
1	0.7 (0.5-1.2)	.26	0.6 (0.4-0.9)	.03	1.0 (0.6-1.7)	.93
2	0.9 (0.5-1.6)	.66	0.9 (0.5-1.8)	.93	1.1 (0.6-2.0)	.69
>3	0.6 (0.3-1.3)	.17	0.7 (0.3-1.5)	.36	1.4 (0.7-3.0)	.38
In-hospital >7 day	N/A	N/A	0.3 (0.1-0.6)	.003	0.5 (0.2-1.0)	.04

^aOR: odds ratio.

^bN/A: not applicable.

^cMedian income has been divided by 10,000 for ease of interpretation.

Figure 4. Forest plot display of the adjusted odds ratios (ORs) and 95% CI for characteristics of response (individual time point and combined models). Forest plot results from all 5 adjusted models are shown. The dashed line indicates an OR of 1. Concurrent mixed modes and race were consistently significant positive predictors of response in all the models. Partnership was also a significant positive predictor of response in most models. Age was a significant negative predictor in the inverse probability weighted model. Postoperative length of stay >7 days was a significant negative predictor in the 7- and 30-day models. Time, which was analyzed in the repeated-measures models, also negatively predicted the response. CMM: concurrent mixed modes; IPW: inverse probability weighted.

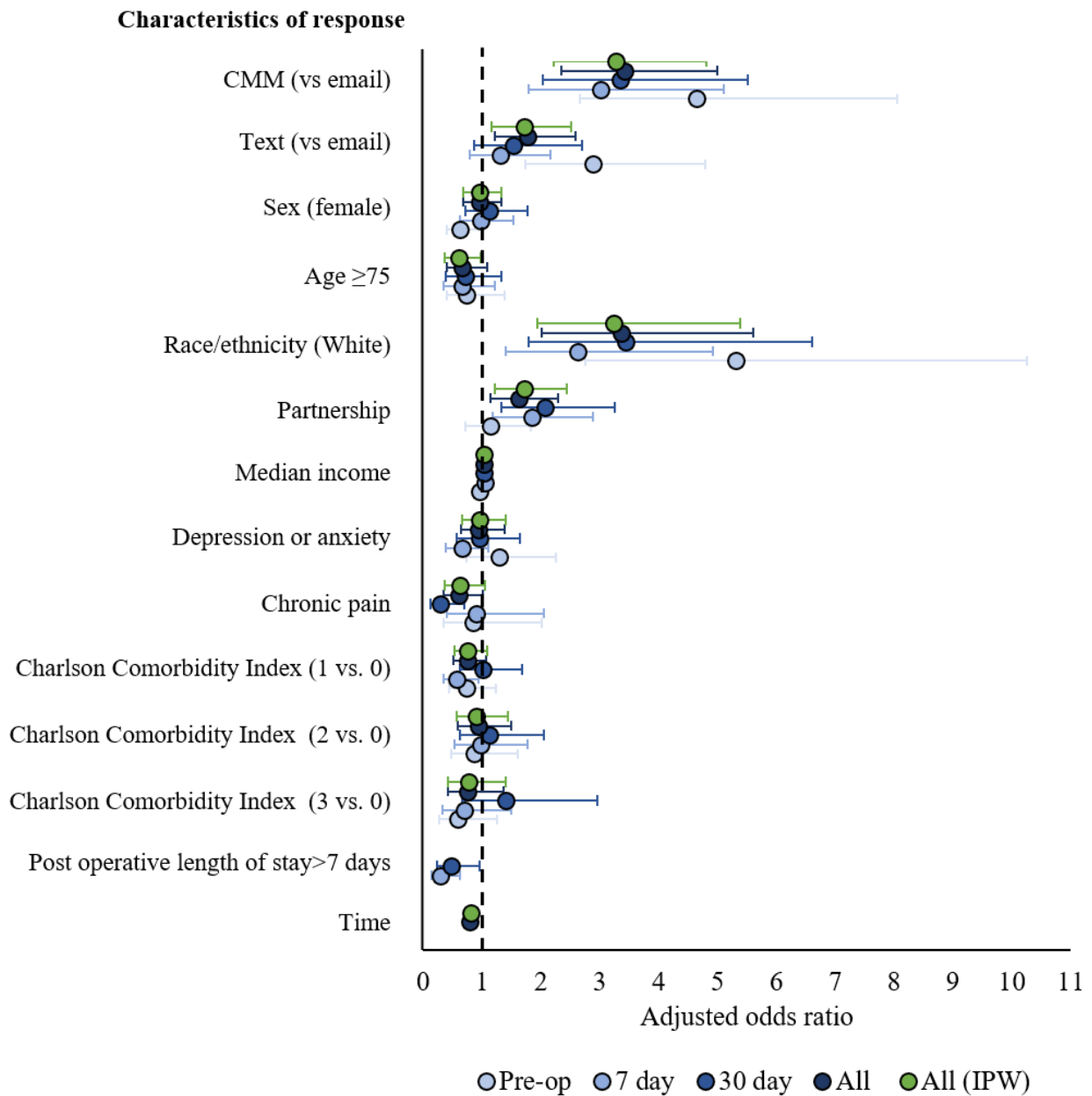


Table 4. Full model generalized estimating equations analysis results for predictors of response.

Characteristics	OR ^a (95% CI)	P value
Method of contact		
Email	Reference	N/A ^b
Text	1.8 (1.2-2.6)	.003
Both	3.4 (2.3-5.0)	<.001
Sex (female)	1.0 (0.7-1.3)	.78
Age ≥75 years	0.7 (0.4-1.1)	.11
Race ethnicity (White)	3.4 (2.0-5.6)	<.001
Partnership	1.6 (1.2-2.3)	.01
Median income ^c	1.0 (0.9-1.1)	.47
Depression or anxiety	0.9 (0.6-1.4)	.75
Chronic pain	0.6 (0.4-1.0)	.06
Charlson comorbidity index		
0	Reference	N/A
1	0.8 (0.5-1.1)	.11
2	0.9 (0.6-1.5)	.78
>3	0.8 (0.4-1.4)	.36
Time	0.8 (0.7-0.9)	<.001

^aOR: odds ratio.

^bN/A: not applicable.

^cMedian income has been divided by 10,000 for ease of interpretation.

Early, Delayed, and Late Responders

There was no evidence of nonresponse bias in the QOR-9 scores (Table 5). 11 surveys were incomplete in the preoperative period and 8 were incomplete in the postoperative period. The surveys were removed from the analysis. Of the remaining 350 patients who responded at the preoperative time point, 212 (60.6%) responded early, 68 (19.4%) responded after the reminder, and 70 (20%) responded via the telephone call. At the 7-day

postoperative time point, 301 participants responded with complete surveys: 171 (56.8%) after the first survey, 79 (26.2%) after the reminder, and 51 (16.9%) via the telephone call. Preoperative and 7-day QOR-9 scores were not significantly different between the early, delayed, and late responders (Table 5). Further sensitivity analyses within subgroups of method of contact, partnership, age, and race demonstrated similar findings (results not shown).

Table 5. Detecting nonresponse bias in outcomes—Quality of Recovery-9 scores for early, delayed, and late responders.

Responder groups	Value, n (%)	Value, median (IQR)	Kruskal–Wallis P value
Preoperative			.73
All	350 (100)	16 (14-17)	
Early responders	212 (60.6)	16 (14-17)	
Delayed responders	68 (19.4)	15 (14-17)	
Late responders	70 (20)	16 (14-17)	
7-day postoperative			.29
All	301 (100)	15 (13-17)	
Early responders	171 (56.8)	15 (13-17)	
Delayed responders	79 (26.2)	15 (13-16)	
Late responders	51 (16.9)	15 (14-17)	

Discussion

Principal Findings

In this one-center enhanced recovery after surgery population, a web-based and mobile communications program appeared to be a feasible approach for collecting longitudinal recovery surveys. Most eligible patients (526/615, 85.5%) opted for the web-based follow-up program, and over half of these patients responded to the survey at each time point. Using CMM to deliver the surveys was an effective strategy to increase response rates compared with email or SMS alone.

However, there were some significant predictors of the inclusion. Those who had a higher number of comorbidities (CCI >1), individuals who were not in a partnership, and individuals over the age of 75 were less likely to opt for the follow-up program. Age was not a significant predictor of response in the adjusted models but became significant after the additional inverse probability weights were applied from the opt-in model. CCI was not a significant predictor of response in any of the adjusted models, including the inverse probability weighted model, whereas partnership was a significant positive predictor in all response models.

Overall response rates declined from the preoperative time point (291/463, 67.1%) to the postoperative time points (256/462, 55.4% and 257/462, 56.6%), but can still be considered relatively high compared with the average response rates for the Hospital Consumer Assessment of Health Care Providers and Systems survey. The Hospital Consumer Assessment of Health Care Providers and Systems survey is a widely used hospital satisfaction survey that uses a combination of mail, telephone, and email to obtain responses from recently discharged patients after surgery and had an average response rate of 23% in New York State in 2018 [26].

The odds of response were significantly higher among those in whom a CMM approach was used to deliver the recovery surveys compared with email alone and SMS alone. These results were consistent over time, even after adjusting for all other risk factors that traditionally predict survey responses. The pattern of CMM achieving the highest response rate over time also remained relatively consistent in different subgroup analyses (eg, within each category of age, race, and partnership). This suggests that using CMM is effective in increasing response rates, even within harder to reach groups. Decomposition analyses confirmed that there were no underlying differences in the observable demographic or clinical characteristics within the 3 method of contact groups that contributed to the differences in response rates, thereby lending credibility to the findings of the response analyses.

There are several possible explanations for why the CMM approach achieved the highest response rates. Giving users a choice to respond with their preferred mode is likely to increase their response [27]. Moreover, receiving the survey prompt to multiple points of contact may have led to an increase in trust. One large survey study in New Zealand explored specific barriers to responses to web-based surveys [28]. Although this survey was conducted in 2009, one of the barriers identified

was trust in relation to spam email or survey requests. Participants associated most survey requests with spam and reported routinely ignoring emails that appeared to be unsolicited at the expense of accidentally ignoring genuine requests. Overall, trust was a minor barrier identified in comparison to issues related to time and effort. Nevertheless, email-related spam is still a major issue, with over half of the current emails estimated to be spam [29]. We do not know the frequency at which our email requests were accidentally and automatically funneled into participants' spam folders, but we reduced this likelihood by using a university email domain. Moreover, by sending the recovery survey to multiple points of contact, we may have increased patients' trust in the legitimacy of the request.

Nonresponse was not associated with lower education or income in the adjusted models. These results are similar to a texting study that assessed pain at 2 weeks postoperatively [11], but contrary to a previous study that sent web-based PRO surveys through email [15]. The latter study consisted of women only but was more nationally representative of the study population. The overall geographic area examined in our study was fairly wealthy and educated, which could explain this difference. The median household income in the United States was approximately US \$63,000, whereas the median household income in our study was approximately US \$97,000.

Race was associated with nonresponse, which is consistent with previous studies [15]. Specifically, those who were Hispanic/non-White had consistently lower odds of responding to the surveys. To understand what effect this might have on interpreting the QOR-9 scores from responders, a separate subgroup analysis was performed to evaluate the risk of nonresponse bias by comparing QOR-9 scores in early, delayed, and late responders who were Hispanic/non-White. Within this subgroup analysis, there were no differences in QOR-9 scores between early, delayed, and late responders at either the preoperative time point or the 7-day time point. However, the sample sizes in these subgroups were small. Nonetheless, these null findings provide modest evidence that even though response rates were lower in this subgroup, those who responded may be representative of the full subgroup.

In fact, no differences were found in the overall and subgroup analyses comparing early, delayed, and late responder recovery scores. Under the continuum of resistance model, nonresponders are most similar to those who require more reminders to complete the survey [30]. Since our recovery scores do not differ between early, delayed, and late respondents, it is possible that nonresponders have similar scores to those included in the sample.

This study has several limitations. First, even though we can describe the patients who did not want to participate (ie, opt-in), we do not know the exact reasons for nonparticipation. It is possible that patients had limited access to technology or that patients simply did not want to participate. We also did not randomize the patients who participated in the 3 method of contact groups. However, decomposition analyses helped alleviate the concern that these groups were somewhat different in composition of characteristics. Indeed, the composition of

the groups was similar, and there did not appear to be trends in characteristics of patients over time (eg, older patients were no more or less likely to have surgery in June compared with November).

Approximately 18% (82/463) of patients received the surveys through a different method of contact other than the one preassigned. This was for various reasons: of the 463 participants, 24 (5.2%) participants in either phase 1 or 3 did not provide an email, and were therefore sent surveys by text alone; 31 (6.7%) participants either in phase 2 or 3 did not own a smartphone, and were therefore sent surveys through email alone; 2 (0.4%) participants in phase 2 did not list a phone number and were sent surveys through email alone; and 25 (5.4%) in phase 2 or 3 either provided contact information that was illegible, or a separate system error occurred in which participants received emails alone. A separate system error was also responsible for shifting several participants into the email and text group at 30 days instead of receiving text or email alone, but this was accounted for in the analyses. Nevertheless, this demonstrates that a system that is preprogrammed to send surveys concurrently through both email and text could prevent the exclusion of individuals who either did not have a smartphone or did not provide an email address by casting a broader net.

Non-White individuals accounted for only 13.3% (82/615) of our sample, most of whom were Hispanic or African Americans. There were too few participants in either the independent race/ethnicity group to perform significance testing separately, but independent examinations did not reveal any inconsistencies between the individual groups and the combined group. The lack of any detected nonresponse bias in recovery scores lends some credence to the representativeness of the sample, but it would still be important to increase response rates among non-White participants and increase the sample size. Furthermore, using late responders as a proxy for nonresponders provides some evidence that nonresponders were similar to responders, but actual scores of nonresponders could not be assessed. In addition, multiple types of surgical procedures were included within this study, which made it impossible to include the procedure as a covariate. Length of stay was used as a proxy

measure for the severity of the surgical procedure and (or) in-hospital complications, but specific surgical area studies could be explored to verify our study findings. Area-level metrics were used as proxies for individual income and education metrics and were much higher than the national averages. These may not truly reflect income and education at the individual level and may also have limited our understanding of how these characteristics affect both inclusion and response.

This study had several strengths. The email or text capture system was applied to the QA program. Therefore, we were able to identify issues of practicality that clinicians might face when attempting to capture postdischarge recovery data in a broad patient population. This provided an opportunity for an in-depth analysis of the generalizability and response rates when using email or text-based programs in surgical patients. At the time, Twilio SMS messages cost US \$0.0075 per message. Therefore, we also demonstrated a resource-friendly method that can obtain high response rates in the early postoperative period. This is important as many complications that occur early after discharge result in readmission [31]. Although we did not actively monitor patient responses, this type of system could in fact be reprogrammed to do so. Although patient surveys contained no identifiable information, they were linked back to a secured database with patient information. Specified algorithms could be developed to send alerts to medical staff regarding patients who appear to be developing a complication or have other unmet medical needs. This would provide an opportunity to intervene before readmission.

Conclusions

A web-based and mobile communication program appears to be a feasible approach for collecting longitudinal perioperative recovery surveys in the ERAS population. Additional efforts may be required to increase participation within non-White individuals, older individuals, and those who are not in a partnership. However, using a CMM approach is an effective strategy to reduce nonresponse, even in difficult-to-reach subgroups. Finally, a web-based system such as the one described could be a cost-effective approach to improve communication between the patient and clinician during a period in which the communication is lacking.

Acknowledgments

The authors would like to thank the institution's Office of the Dean of Medicine and the Department of Institutional Technology (specifically Tony Jin), as well as Gemma Aguilar, RN, at the Pre-Operative Services Clinic for support during this project.

Conflicts of Interest

None declared.

References

1. Jiang H, Jian X, Shanguan Y, Qing J, Chen L. Effects of enhanced recovery after surgery in total knee arthroplasty for patients older than 65 years. *Orthop Surg* 2019 Apr;11(2):229-235 [FREE Full text] [doi: [10.1111/os.12441](https://doi.org/10.1111/os.12441)] [Medline: [30945802](https://pubmed.ncbi.nlm.nih.gov/30945802/)]
2. Semerjian A, Milbar N, Kates M, Gorin MA, Patel HD, Chalfin HJ, et al. Hospital charges and length of stay following radical cystectomy in the enhanced recovery after surgery era. *Urology* 2018 Jan;111:86-91. [doi: [10.1016/j.urology.2017.09.010](https://doi.org/10.1016/j.urology.2017.09.010)] [Medline: [29032237](https://pubmed.ncbi.nlm.nih.gov/29032237/)]

3. Baek H, Cho M, Kim S, Hwang H, Song M, Yoo S. Analysis of length of hospital stay using electronic health records: a statistical and data mining approach. *PLoS One* 2018;13(4):e0195901 [FREE Full text] [doi: [10.1371/journal.pone.0195901](https://doi.org/10.1371/journal.pone.0195901)] [Medline: [29652932](https://pubmed.ncbi.nlm.nih.gov/29652932/)]
4. Stessel B, Fiddelaers AA, Joosten EA, Hoofwijk DM, Gramke H, Buhre WF. Prevalence and predictors of quality of recovery at home after day surgery. *Medicine (Baltimore)* 2015 Sep;94(39):e1553 [FREE Full text] [doi: [10.1097/MD.0000000000001553](https://doi.org/10.1097/MD.0000000000001553)] [Medline: [26426622](https://pubmed.ncbi.nlm.nih.gov/26426622/)]
5. Bowyer A, Royse CF. The future of postoperative quality of recovery assessment: multidimensional, dichotomous, and directed to individualize care to patients after surgery. *Curr Opin Anaesthesiol* 2016 Dec;29(6):683-690. [doi: [10.1097/ACO.0000000000000399](https://doi.org/10.1097/ACO.0000000000000399)] [Medline: [27764047](https://pubmed.ncbi.nlm.nih.gov/27764047/)]
6. Bowyer AJ, Royse CF. Postoperative recovery and outcomes--what are we measuring and for whom? *Anaesthesia* 2016 Jan;71 Suppl 1:72-77 [FREE Full text] [doi: [10.1111/anae.13312](https://doi.org/10.1111/anae.13312)] [Medline: [26620150](https://pubmed.ncbi.nlm.nih.gov/26620150/)]
7. Basch E. Patient-reported outcomes - harnessing patients' voices to improve clinical care. *N Engl J Med* 2017 Jan 12;376(2):105-108. [doi: [10.1056/NEJMp1611252](https://doi.org/10.1056/NEJMp1611252)] [Medline: [28076708](https://pubmed.ncbi.nlm.nih.gov/28076708/)]
8. Abola RE, Bennett-Guerrero E, Kent ML, Feldman LS, Fiore JF, Shaw AD, Perioperative Quality Initiative (POQI) 2 Workgroup. American society for enhanced recovery and perioperative quality initiative joint consensus statement on patient-reported outcomes in an enhanced recovery pathway. *Anesth Analg* 2018 Jun;126(6):1874-1882. [doi: [10.1213/ANE.0000000000002758](https://doi.org/10.1213/ANE.0000000000002758)] [Medline: [29293180](https://pubmed.ncbi.nlm.nih.gov/29293180/)]
9. Mobile fact sheet. Pew Research Center. 2019. URL: <https://www.pewinternet.org/fact-sheet/mobile/> [accessed 2019-06-01]
10. Carrier G, Cotte E, Beyer-Berjot L, Faucheron JL, Joris J, Slim K, Groupe Francophone de Réhabilitation Améliorée après Chirurgie (GRACE). Post-discharge follow-up using text messaging within an enhanced recovery program after colorectal surgery. *J Visc Surg* 2016 Aug;153(4):249-252. [doi: [10.1016/j.jviscsurg.2016.05.016](https://doi.org/10.1016/j.jviscsurg.2016.05.016)] [Medline: [27423211](https://pubmed.ncbi.nlm.nih.gov/27423211/)]
11. Anthony CA, Lawler EA, Ward CM, Lin IC, Shah AS. Use of an automated mobile phone messaging robot in postoperative patient monitoring. *Telemed J E Health* 2018 Jan;24(1):61-66. [doi: [10.1089/tmj.2017.0055](https://doi.org/10.1089/tmj.2017.0055)] [Medline: [28622079](https://pubmed.ncbi.nlm.nih.gov/28622079/)]
12. Anthony CA, Volkmar A, Shah AS, Willey M, Karam M, Marsh JL. Communication with orthopedic trauma patients via an automated mobile phone messaging robot. *Telemed J E Health* 2018 Jul;24(7):504-509. [doi: [10.1089/tmj.2017.0188](https://doi.org/10.1089/tmj.2017.0188)] [Medline: [29261036](https://pubmed.ncbi.nlm.nih.gov/29261036/)]
13. Nelson SE, Adams AJ, Buczek MJ, Anthony CA, Shah AS. Postoperative pain and opioid use in children with supracondylar humeral fractures: balancing analgesia and opioid stewardship. *J Bone Joint Surg Am* 2019 Jan 16;101(2):119-126. [doi: [10.2106/JBJS.18.00657](https://doi.org/10.2106/JBJS.18.00657)] [Medline: [30653041](https://pubmed.ncbi.nlm.nih.gov/30653041/)]
14. Lucas SM, Kim T, Ghani KR, Miller DC, Linsell S, Starr J, Michigan Urological Surgery Improvement Collaborative. Establishment of a web-based system for collection of patient-reported outcomes after radical prostatectomy in a statewide quality improvement collaborative. *Urology* 2017 Sep;107:96-102. [doi: [10.1016/j.urology.2017.04.058](https://doi.org/10.1016/j.urology.2017.04.058)] [Medline: [28652160](https://pubmed.ncbi.nlm.nih.gov/28652160/)]
15. Berlin NL, Hamill JB, Qi J, Kim HM, Pusic AL, Wilkins EG. Nonresponse bias in survey research: lessons from a prospective study of breast reconstruction. *J Surg Res* 2018 Apr;224:112-120 [FREE Full text] [doi: [10.1016/j.jss.2017.11.058](https://doi.org/10.1016/j.jss.2017.11.058)] [Medline: [29506826](https://pubmed.ncbi.nlm.nih.gov/29506826/)]
16. Myles PS. More than just morbidity and mortality - quality of recovery and long-term functional recovery after surgery. *Anaesthesia* 2020 Jan;75 Suppl 1:143-150 [FREE Full text] [doi: [10.1111/anae.14786](https://doi.org/10.1111/anae.14786)] [Medline: [31903564](https://pubmed.ncbi.nlm.nih.gov/31903564/)]
17. Myles PS, Hunt JO, Nightingale CE, Fletcher H, Beh T, Tanil D, et al. Development and psychometric testing of a quality of recovery score after general anesthesia and surgery in adults. *Anesth Analg* 1999 Jan;88(1):83-90. [doi: [10.1097/00000539-199901000-00016](https://doi.org/10.1097/00000539-199901000-00016)] [Medline: [9895071](https://pubmed.ncbi.nlm.nih.gov/9895071/)]
18. Anderson M, Perrin A. Tech adoption climbs among older adults: technology use among seniors. Pew Research Center. 2017. URL: <https://www.pewinternet.org/2017/05/17/technology-use-among-seniors/> [accessed 2019-06-01]
19. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-383. [doi: [10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)] [Medline: [3558716](https://pubmed.ncbi.nlm.nih.gov/3558716/)]
20. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res* 2014 Oct;23(5):440-459. [doi: [10.1177/0962280213476378](https://doi.org/10.1177/0962280213476378)] [Medline: [23427225](https://pubmed.ncbi.nlm.nih.gov/23427225/)]
21. Bell ML, Horton NJ, Dhillion HM, Bray VJ, Vardy J. Using generalized estimating equations and extensions in randomized trials with missing longitudinal patient reported outcome data. *Psychooncology* 2018 Sep;27(9):2125-2131. [doi: [10.1002/pon.4777](https://doi.org/10.1002/pon.4777)] [Medline: [29802657](https://pubmed.ncbi.nlm.nih.gov/29802657/)]
22. Bell ML, Floden L, Rabe BA, Hudgens S, Dhillion HM, Bray VJ, et al. Analytical approaches and estimands to take account of missing patient-reported data in longitudinal studies. *Patient Relat Outcome Meas* 2019;10:129-140 [FREE Full text] [doi: [10.2147/PROM.S178963](https://doi.org/10.2147/PROM.S178963)] [Medline: [31114411](https://pubmed.ncbi.nlm.nih.gov/31114411/)]
23. Lin IF, Schaeffer NC. Using survey participants to estimate the impact of nonparticipation. *Pub Opin Q* 1995;59(2):236-258. [doi: [10.1086/269471](https://doi.org/10.1086/269471)]
24. Cella D, Han E, Jensen S. Patient-reported outcomes in performance measurement. *Res Triangle Inst* 2015:978-971. [doi: [10.3768/rtipress.2015.bk.0014.1509](https://doi.org/10.3768/rtipress.2015.bk.0014.1509)] [Medline: [28211667](https://pubmed.ncbi.nlm.nih.gov/28211667/)]
25. Thomson M, Pledger M, Hamblin R, Cumming J, Tawfiq E. Comparing initial and follow-up responders to a New Zealand patient experience survey. *N Z Med J* 2018 Sep 21;131(1482):46-58. [Medline: [30235192](https://pubmed.ncbi.nlm.nih.gov/30235192/)]

26. CAHPS Hospital Survey. Centers for Medicare & Medicaid Services. 2020. URL: <https://hcahpsonline.org/en/summary-analyses/> [accessed 2020-06-24]
27. Dillman D, Smyth J, Christian L. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method. Fourth ed. Hoboken, New Jersey: John Wiley & Sons, Inc; 2014.
28. Han V, Albaum G, Wiley JB, Thirkell P. Applying theory to structure respondents' stated motivations for participating in web surveys. *Qual Mrkt Res: An Int J* 2009 Sep 04;12(4):428-442. [doi: [10.1108/13522750910993338](https://doi.org/10.1108/13522750910993338)]
29. Clement J. Global spam volume as percentage of total e-mail traffic from January 2014 to March 2020, by month. Statista. 2020. URL: <https://www.statista.com/statistics/420391/spam-email-traffic-share/> [accessed 2020-10-15]
30. Boniface S, Scholes S, Shelton N, Connor J. Assessment of non-response bias in estimates of alcohol consumption: applying the continuum of resistance model in a general population survey in England. *PLoS One* 2017;12(1):e0170892 [FREE Full text] [doi: [10.1371/journal.pone.0170892](https://doi.org/10.1371/journal.pone.0170892)] [Medline: [28141834](https://pubmed.ncbi.nlm.nih.gov/28141834/)]
31. Tevis SE, Kohlhofer BM, Weber SM, Kennedy GD. Postdischarge complications are an important predictor of postoperative readmissions. *Am J Surg* 2014 Oct;208(4):505-510 [FREE Full text] [doi: [10.1016/j.amjsurg.2014.05.013](https://doi.org/10.1016/j.amjsurg.2014.05.013)] [Medline: [25150195](https://pubmed.ncbi.nlm.nih.gov/25150195/)]

Abbreviations

CCI: Charlson Comorbidity Index
CMM: concurrent mixed modes
ERAS: enhanced recovery after surgery
IPW: inverse probability weighted
OR: odds ratio
PRO: patient-reported outcomes
QA: quality assurance
QOR-9: Quality of Recovery-9

Edited by G Eysenbach; submitted 22.10.20; peer-reviewed by I Hochberg, B Smith; comments to author 29.01.21; revised version received 25.02.21; accepted 23.09.21; published 04.11.21.

Please cite as:

Romeiser JL, Cavalcante J, Richman DC, Singh SM, Liang X, Pei A, Sharma S, Lazarus Z, Gan TJ, Bennett-Guerrero E
Comparing Email, SMS, and Concurrent Mixed Modes Approaches to Capture Quality of Recovery in the Perioperative Period: Retrospective Longitudinal Cohort Study
JMIR Form Res 2021;5(11):e25209
URL: <https://formative.jmir.org/2021/11/e25209>
doi: [10.2196/25209](https://doi.org/10.2196/25209)
PMID: [34734827](https://pubmed.ncbi.nlm.nih.gov/34734827/)

©Jamie L Romeiser, James Cavalcante, Deborah C Richman, Sunitha M Singh, Xiaohui Liang, Allison Pei, Samanvaya Sharma, Zoe Lazarus, Tong J Gan, Elliott Bennett-Guerrero. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 04.11.2021. 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 Virtual Community for Disability Advocacy: Development of a Searchable Artificial Intelligence–Supported Platform

Christo El Morr¹, PhD; Pierre Maret², PhD; Fabrice Muhlenbach², PhD; Dhayananth Dharmalingam², MSc; Rediet Tadesse², MSc; Alexandra Creighton¹, MA; Bushra Kundi¹, BHS; Alexis Buettgen¹, PhD; Thumeka Mgwigwi³, MA; Serban Dinca-Panaitescu¹, PhD; Enakshi Dua⁴, PhD; Rachel Gorman¹, PhD

¹School of Health Policy and Management, Faculty of Health, York University, Toronto, ON, Canada

²CNRS, UMR 5516, Laboratoire Hubert Curien, Université Jean Monnet Saint Etienne, Saint Etienne, France

³Student Learning and Academic Success Department, York University Libraries, York University, Toronto, ON, Canada

⁴School of Gender, Sexuality and Women's Studies, York University, Toronto, ON, Canada

Corresponding Author:

Christo El Morr, PhD

School of Health Policy and Management

Faculty of Health

York University

Stong College Room 306

4700 Keele St

Toronto, ON, M3J 1P3

Canada

Phone: 1 4167362100

Email: elmorr@yorku.ca

Abstract

Background: The lack of availability of disability data has been identified as a major challenge hindering continuous disability equity monitoring. It is important to develop a platform that enables searching for disability data to expose systemic discrimination and social exclusion, which increase vulnerability to inequitable social conditions.

Objective: Our project aims to create an accessible and multilingual pilot disability website that structures and integrates data about people with disabilities and provides data for national and international disability advocacy communities. The platform will be endowed with a document upload function with hybrid (automated and manual) paragraph tagging, while the querying function will involve an intelligent natural language search in the supported languages.

Methods: We have designed and implemented a virtual community platform using Wikibase, Semantic Web, machine learning, and web programming tools to enable disability communities to upload and search for disability documents. The platform data model is based on an ontology we have designed following the United Nations Convention on the Rights of Persons with Disabilities (CRPD). The virtual community facilitates the uploading and sharing of validated information, and supports disability rights advocacy by enabling dissemination of knowledge.

Results: Using health informatics and artificial intelligence techniques (namely Semantic Web, machine learning, and natural language processing techniques), we were able to develop a pilot virtual community that supports disability rights advocacy by facilitating uploading, sharing, and accessing disability data. The system consists of a website on top of a Wikibase (a Semantic Web–based datastore). The virtual community accepts 4 types of users: information producers, information consumers, validators, and administrators. The virtual community enables the uploading of documents, semiautomatic tagging of their paragraphs with meaningful keywords, and validation of the process before uploading the data to the disability Wikibase. Once uploaded, public users (information consumers) can perform a semantic search using an intelligent and multilingual search engine (QAnswer). Further enhancements of the platform are planned.

Conclusions: The platform ontology is flexible and can accommodate advocacy reports and disability policy and legislation from specific jurisdictions, which can be accessed in relation to the CRPD articles. The platform ontology can be expanded to fit international contexts. The virtual community supports information upload and search. Semiautomatic tagging and intelligent multilingual semantic search using natural language are enabled using artificial intelligence techniques, namely Semantic Web, machine learning, and natural language processing.

KEYWORDS

virtual community; machine learning; Semantic Web; natural language processing; web intelligence; health informatics; Wikibase; disability rights; human rights; CRPD; equity; community; disability; ethics; rights; pilot; platform

Introduction

Human rights monitoring for people with disabilities is a pressing issue in Canada and France, as well as internationally. However, data and information about human rights policy and legal precedents, lived experiences, and media portrayals of people with disabilities are scarce. This limited data and information hinders the capacity for knowledge mobilization among researchers, nongovernmental organizations, and advocacy groups in their efforts to monitor the implementation and realization of human rights for people with disabilities. In addition, policy-based definitions of disability vary widely across local, national, and international jurisdictions, creating further challenges regarding the access of comparative data and information.

Internationally, a few sources provide access to some disability data (eg, the Disability Data Advocacy Working Group [1], Disability Data Portal [2], and Disability Rights Promotion International [DRPI] [3]). However, these databases have limited data while the need for such data is overwhelming. The incorporation of disability into human rights–focused legislations, constitutions, and international law (eg, the United Nations Convention on the Rights of Persons with Disabilities, or CRPD [4]) has highlighted an ongoing need for access to disability data for disability researchers, organizations of people with disabilities, and broader public policy and development sectors. The lack of availability of disability data has been identified as a major challenge for continuous disability rights monitoring (ie, collection, integration, analysis, interpretation, and mobilization of data and knowledge about the implementation and realization of human rights for people with disabilities) [5,6]. It is important to develop an inclusive approach to analytics and machine learning, one that can be scaled up to integrate disability data and enable the searching of such data to expose systemic discrimination and social exclusion, which increase vulnerability to inequitable social conditions [7].

Historically, there has been a lack of data on disability collected and shared on a global and national level [8]. The lack of aggregate and disaggregate disability data available is hampering researchers' ability to conduct critical health research [9]. It is also limiting our knowledge about the experiences of people with disabilities and the challenges and barriers they face [6]. Reed, Meeks, and Swenor [10] argue that “the disability data gap is more than just a surveillance oversight; social injustices exist that cannot be separated from this lack of information. The lack of data perpetuates the exclusion of disabled people from discussions of health equity and policies that are data driven.”

Disability is a complex multidimensional experience that can be difficult to measure. Operational measures of disability vary

according to the purpose and application of the data, the conception of disability, and the aspects of disability examined [10]. However, access to disability data is an issue of human rights. For example, Article 31 of the CRPD [11] notes the following obligation:

States Parties undertake to collect appropriate information, including statistical and research data, to enable them to formulate and implement policies to give effect to the present Convention... The information collected shall be disaggregated, as appropriate, and... States Parties shall assume responsibility for the dissemination of these statistics and ensure their accessibility to persons with disabilities and others.

The disability Wikibase addresses this responsibility by making data accessible and available to a variety of stakeholders. Understanding, interpreting, and using disability data is essential to stakeholders with interests in disability issues, including advocacy groups, individuals, and organizations of people with disabilities, nongovernmental organizations, policy makers, and researchers [12]. Data contribute to decisions about poverty alleviation strategies, health and social service development, and humanitarian responses. Data collection is critical to the formulation of evidence-based policies and all aspects of the implementation of disability-inclusive policies and programs [13]. Reliable, accessible data are also crucial for monitoring progress and implementation of the CRPD.

Meanwhile, advancements are being made in the field of health informatics, which can assist in the organization and use of health data. Health informatics is an interdisciplinary field that uses computer science, decision-making, information processing, and management science to design and implement information systems in a user-centered manner to get, store, process, and provide information for users to assist them in achieving their objectives [14]. Some of the techniques used in health informatics are artificial intelligence (AI) techniques [15]. Machine learning techniques are trained algorithms that assist users in making sense of data to make organizational decisions [16] and health decisions [17,18]. Semantic Web techniques and natural language processing (NLP) provide computers with the ability to decipher the meaning of information and process phrases written or spoken in natural language [19]. The combination of these techniques can allow users to search for data using natural language sentences [20,21] in multiple languages, and they can also be used to find semantic similarities among disparate documents from different disciplines [22–26].

On the other hand, techniques to store, integrate, and search for data do exist; the publicly available Wikibase environment is an open-source collaborative tool for editing, integrating, and storing structured data. Wikidata is the most famous Wikibase example, and it targets public knowledge [27].

Our project aims to create a screen reader-accessible multilingual website on top of a Wikibase that structures and integrates data about people with disabilities and serves national and international disability advocacy communities. This pilot website will be endowed with a document upload function with an automated data-from-text extraction function for insertion into the Wikibase, and with an intelligent natural language search capability for querying the data.

Methods

Choice of Domain and Documents

To design and develop the underlying data model of the pilot platform, our domain of choice was the CRPD. The CRPD is an international human rights treaty intended to articulate and protect the rights and dignity of people with disabilities. The CRPD reflects a shift from viewing people with disabilities as objects of charity, medical treatment, and social protection toward viewing them as full and equal members of society with human rights. The CRPD contains 50 Articles covering a range of issues including health, accessibility, data collection, and monitoring mechanisms and is informed by the principles of dignity, autonomy, nondiscrimination, equality, and respect for difference. The CRPD has been signed by over 150 countries, including Canada and France.

This domain of choice allows for a large number of documents connected to the CRPD to be uploaded in the future, enabling the testing of the platform on a large scale. Our first set of documents to be shared included DRPI's reports of country- and region-specific rights monitoring based on the CRPD. DRPI is a collaborative project to establish a comprehensive, sustainable international system to monitor the human rights of people with disabilities in accordance with the CRPD. This project has been operating for over a decade and has collected data and information from North America, Europe, Latin America, Asia Pacific, Africa, the Middle East, and North Africa.

Platform Design

In early stages of the project, we collected the requirements of the platform and identified that the platform should act as a virtual community and incorporate the following characteristics:

1. Facilitate knowledge sharing by enabling the uploading of disability documents.
2. Provide validated information by allowing only authenticated external partners (eg, community organizations) to add documents and having a user approve them, acting as a gatekeeper.
3. Support disability rights advocacy by enabling dissemination of knowledge to the public by making the documents available to all potential users and using a natural language-enabled search engine.
4. Administer the platform to create users, continuously refine the machine learning models, provide reports, etc.

In summary, the aim of this project was to create a virtual knowledge network using the Semantic Web and other AI techniques. The virtual knowledge network includes 4 types of users:

1. Authenticated information producers that upload information of interest (eg, reports, blogs, articles) and assign tags (ie, keywords) to the new data.
2. Domain expert users that play the role of gatekeepers and validate the uploaded information (ie, documents and tags) by issuing an upload request to the platform administrator.
3. Platform administrators that manage the platform. Their tasks include approving/rejecting upload requests, continuously maintaining the platform, and finetuning the AI models.
4. Information consumers that search for information stored on the platform using natural language questions.

Platform Implementation

The virtual knowledge network front end includes a simple website connected to a search engine, while the back end uses the following:

1. A machine learning model to tag each paragraph of the uploaded documents with corresponding keywords that reflect the meanings embedded in the text.
2. A Wikibase structure that holds the disability rights knowledge base and the corresponding data (ie, tags, annotated paragraphs, links to the documents).
3. An intelligent search engine (QAnswer) that uses NLP to search the Wikibase for answers to users' questions.

To implement the different parts of the platform, the following major programming stacks were used:

1. React JS, a simple and flexible JavaScript-based library, was used to develop the front end of the platform [28].
2. MySQL was our choice for the database used to store temporary document information before its semantification and storage in the Wikibase [29].
3. Wikibase was used to store the disability rights knowledge base, which was represented in a knowledge graph and used the Resource Description Framework (RDF) format [30]. A knowledge graph is a network of entities, their semantic types and properties, and the relationships between the entities [31]. It is used to represent information and reasoning.
4. QAnswer, a platform that makes RDF data accessible via natural language, has been used to enable intelligent search using intelligent natural language questioning [32-34]. QAnswer is the first AI-driven platform to query RDF datastores in natural language. It uses semantic technologies (NLP, Semantic Web) as well as machine learning techniques [32].
5. Python was used to develop cross-platform and server-side applications [35].
6. FastText was used for NLP to create a machine learning model to tag paragraphs. FastText is often on par with deep learning classifiers in terms of accuracy, and is many orders of magnitude faster for training and evaluation [36,37].
7. PyMuPDF, a Python library, was used to access and process uploaded documents [38].
8. GitHub was used to maintain source code versioning [39].

Figure 2. A partial knowledge graph that depicts entities and relationships present in CRPD Article 25. CRPD: Convention on the Rights of Persons with Disabilities; DRPI: Disability Rights Promotion International.

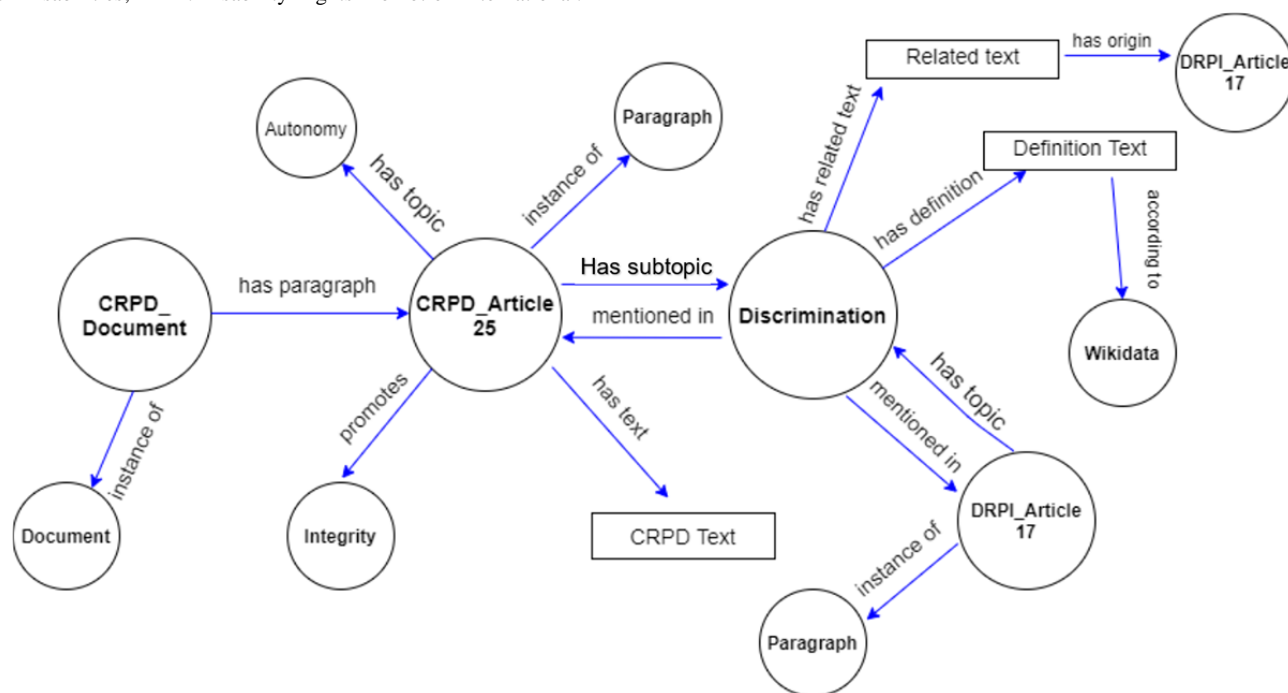


Figure 2 shows the relationships among the entities of articles 25 and 17 of the CRPD using a knowledge graph. Documents, topics, and sources are concepts connected by relations (eg, “has topic,” “has paragraph,” “according to,” “mentioned in,” “has text,” “has definition,” “has origin,” “instance of”). For instance, paragraphs are connected to the document using the “has paragraph” relation since each document is formed of many paragraphs. In turn, each paragraph has one or more topics and subtopics, and so on.

Topic Modeling Using Latent Dirichlet Allocation

We used a latent Dirichlet allocation (LDA) [40] machine learning algorithm to model the topics in the available documents. To train the model, a set of 30 unstructured and unlabeled documents were provided in an unsupervised manner. To determine the number of topics to include in this unsupervised learning method, we conducted tests to select the best coherence measure score; this score is based on the degree of semantic similarity between high-scoring words in the topic.

The LDA result provided disability experts with an initial list of keywords that could be used to tag the documents’ content. The disability experts added to the list of keywords and created a glossary of terms related to the disability field. The glossary terms were used as labels for tagging the paragraphs of a training set of the document corpus. Afterward, the labeled documents

were used by the supervised machine learning algorithm proposed by the FastText library to design a text classification model that allows users to tag uploaded documents.

Initial Glossary Development and Enrichment

Disability domain experts developed a glossary of keywords that describe the main topics and subtopics embedded in the DRPI and CRPD documents. These keywords were then used to tag paragraphs of the uploaded documents with appropriate topics.

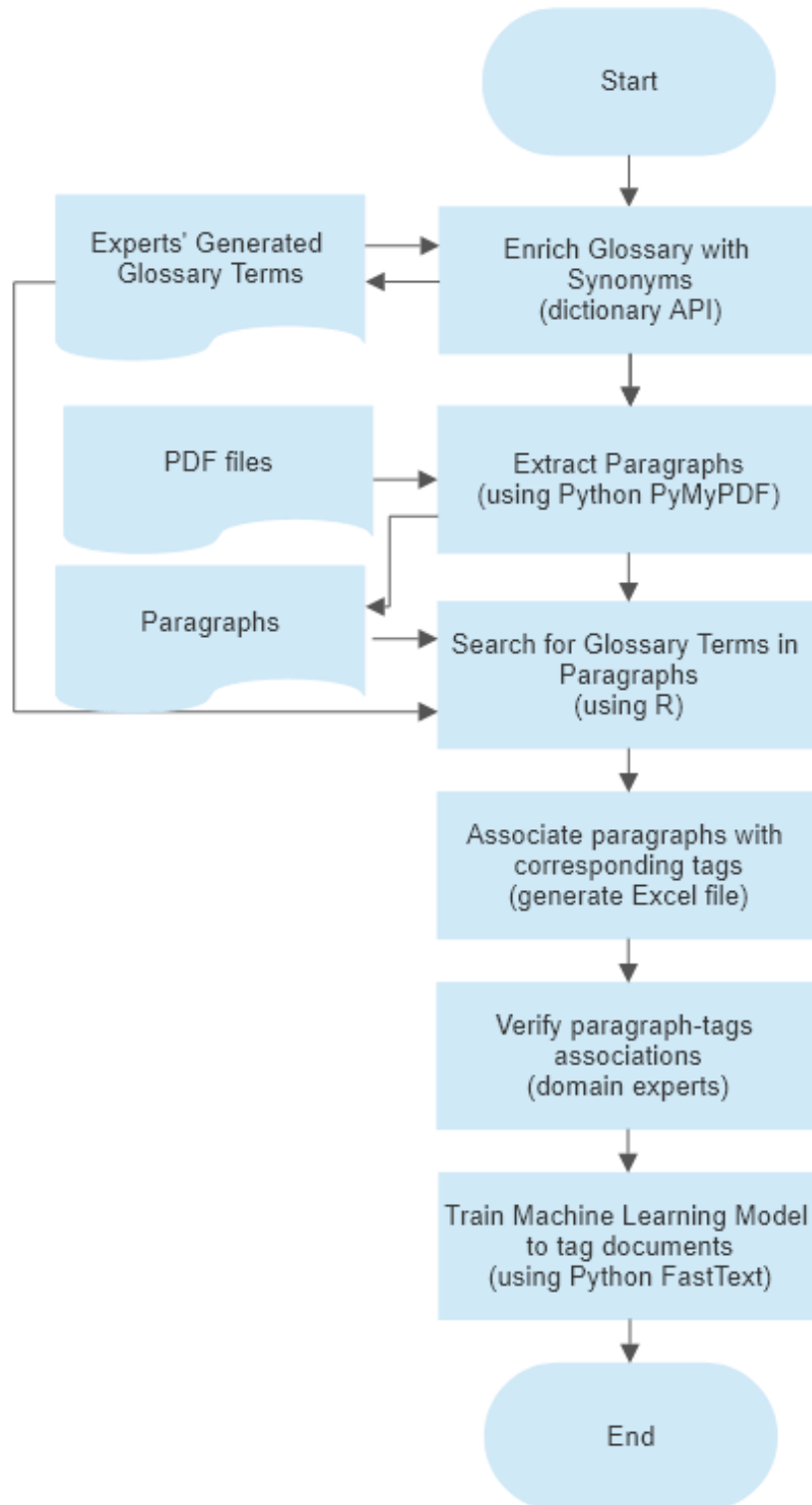
Following the development of the glossary, the Merriam-Webster Dictionary application programming interface (API) was used to get synonyms of the glossary terms. Thus, the glossary terms were enriched with synonym words.

Document Tagging

Overview

Once a user uploads a document, the platform splits the document into paragraphs and uses the glossary terms to tag them with appropriate topics. In the first phase, the extracted paragraphs and corresponding tags were verified by domain experts and the result was used as a training data set for a machine learning algorithm (Figure 3). The algorithm was used on the platform to automatically tag newly uploaded documents.

Figure 3. Document tagging flowchart. API: application programming interface.



Paragraph Extraction and Tagging

Using the PyMuPDF library (a Python binding for MuPDF), the PDF files were parsed, and paragraphs were extracted. Afterward, R (version 4.0.5; R Foundation for Statistical Computing) was used to tag the paragraphs with appropriate glossary terms. The tags were reviewed by domain experts who added or deleted tags as needed. The result was a set of validated

instances of paragraphs and their corresponding tags, which were used as the training data set.

Machine Learning Model Generation to Automate Tagging

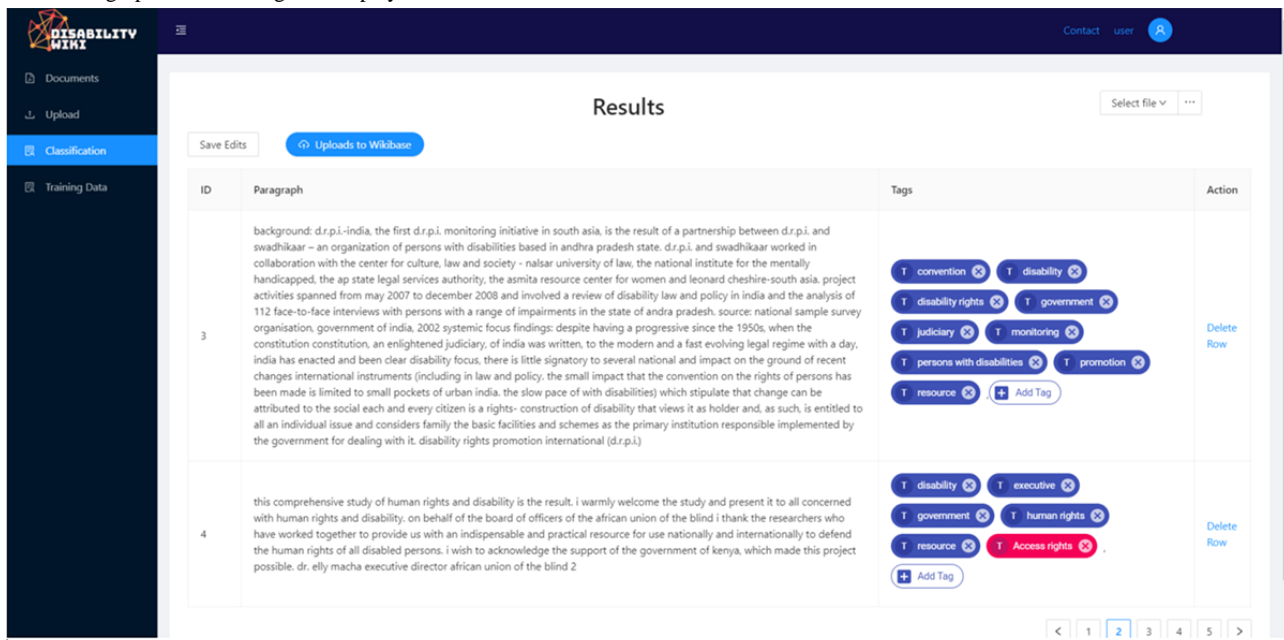
The validated paragraph-tag instances were used as a training set for the FastText algorithm. FastText is a library for efficient learning of word representations and sentence classification. This library can be used to build a word representation model

with unsupervised training (using SkipGram or CBOW), or to design a text classification model with supervised learning. We used the latter approach and used the FastText library defaults for tagging the document paragraphs.

The FastText algorithm proposes tags from the glossary for the paragraphs and the result is displayed (Figure 4). Users can

reject or add to the proposed tags, allowing for continuous enrichment of the glossary and training of the model, which make it more accurate. A request to upload the tagged paragraphs and other related information to the Wikibase is sent to the administrator.

Figure 4. Paragraphs and their tags are displayed to the user.

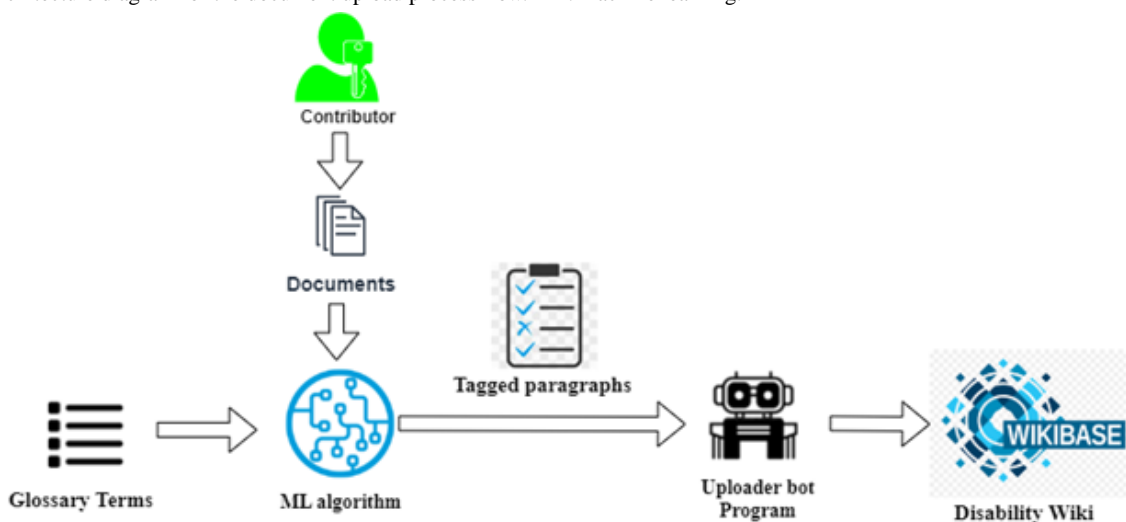


Document Upload to Wikidata

The administrator can review the upload requests and accept or reject them. Upon acceptance, the tagged paragraphs are

semantified based on the ontology and uploaded to the disability Wikibase, where they are stored as triples that can be subject to semantic search (Figure 5).

Figure 5. Architecture diagram for the document upload process flow. ML: machine learning.

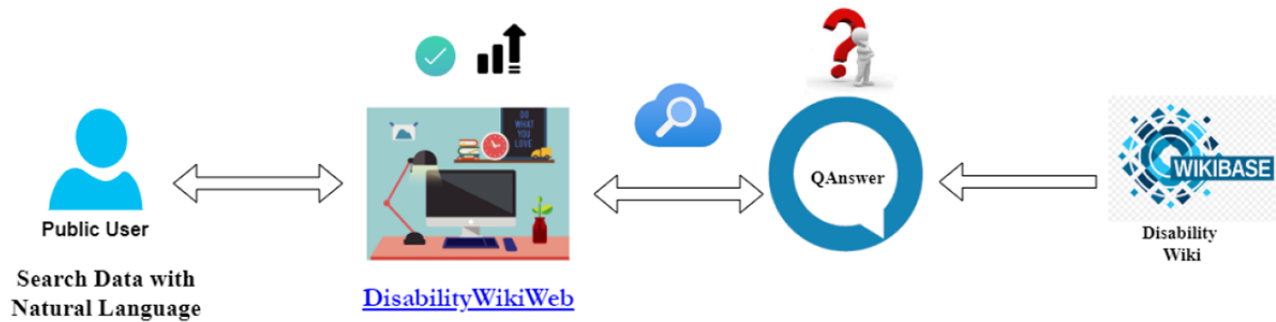


Natural Language Search Using QAnswer

The disability Wikibase is connected to the QAnswer search engine, which performs a natural language search (Figure 6). A simple search text field on the website is provided for users to enter their queries.

Although only authorized users can be producers and upload documents, any user can be a consumer and search for information. The disability Wikibase uses QAnswer API endpoints to send questions to, and retrieve answers from, QAnswer. We used the autocomplete feature in QAnswer to provide several precompleted questions while users are typing their questions. Thus, users can select from these autocompleted options instead of continuing to type their question.

Figure 6. Search function flow of the disability wiki website.



Discussion

Principal Results

There are currently no tools available to collect and share data related to people with disabilities. We have developed a platform that can fill this gap and respond to this need.

Using health informatics and AI techniques, we were able to develop a pilot virtual community that supports disability rights advocacy by facilitating the sharing and accessing of disability data. The virtual community accepts 4 types of users: information producers, information consumers, validators, and administrators. The virtual community enables the uploading of documents, semiautomatic tagging of document paragraphs with meaningful keywords, and validation of the process before the document and its associated tags are uploaded to the disability Wikibase. Once uploaded, public users can perform a semantic search using the QAnswer intelligent search engine. Currently, the platform allows information producers to upload documents and label their paragraphs with appropriate tags, and also allows information consumers to search for information using natural language. Hence, the platform answers a serious need in the disability community to store disability-related data in one database that includes an intelligent search function.

Virtual communities for health purposes (including disabilities) already exist; however, their purpose is to disseminate information only, with no means for members to contribute additional data and information. Wikipedia is wiki-based but it is not specialized and includes nonvalidated, general-purpose information, with no means to upload documents or perform a semantic search. This is the first time an AI approach was used to bolster disability advocacy. The proposed knowledge graph's content has been validated by disability domain experts, and it is also technically sound—Wikidata's best practices were followed. Future evolutions of the knowledge graph will have to maintain its accuracy, consistency, and conciseness, as defined by Zaveri et al [41].

This project advances our previous informatics work on a virtual knowledge network using documents from the DRPI [42]; in our previous work, we developed a virtual knowledge network to share documents [42], while on this pilot site we enabled semantic search and allowed accurate tagging of documents' content using machine learning approaches. The current platform simplifies document sharing and offers the public, disability organizations, researchers, and policy makers a tool to allow for more transparency in disability data.

Our proposed generic novel approach is to provide a platform for semantic free-text search on freely collected documents on a given topic. Our future research aims to answer the following question: what kind of information is searched for and by what types of users? To answer this question, we are planning to collect information regarding the users of the platform and the search terms they use; in addition, a thematic analysis will also be conducted to understand the users' needs.

Limitations and Future Perspectives

This work is still in its pilot stage; there is a limited number of documents available for machine learning training, while accuracy measurement and evaluation of the machine learning algorithms are underway. However, the more users upload documents and validate the automatic tagging, the more accurate the automatic tagging becomes due to continuous retraining.

There are other features and project phases planned for the future. Virtual communities have been used in many domains, including health care [43], to address chronic disease management, education [44], and mental health. Our current platform is expanding the use of virtual communities for social purposes in the domain of disability rights [45]. The next phases of the project will involve an evaluation of the adoption of this virtual community using evaluation frameworks [46], as well as the use of background analytics to analyze usage patterns and enhance the search capacity. Moreover, users' experience will be formally evaluated to enhance acceptance and adherence. Universal accessibility allows equitable access to information; accessibility testing is underway, and improvements are being implemented. Additionally, while QAnswer can answer questions formulated in English and French, the current platform is written in English; translating the website into French and uploading documents written in French are both on our project agenda.

NLP has been used for many health care applications for information extraction, sentiment analysis, search extraction, and information synthesis [47], using the state-of-the-art NLP model Bidirectional Encoder Representations from Transformers (BERT) [48]. Our next steps include tagging optimization using newer approaches to NLP such as BERT. This platform may encourage the creation of a domain-specific DisabilityBERT, similar to the BioBERT adapted for biomedical text and SciBERT adapted for science text [49].

In our future work, we will continue to elucidate biases within the disability Wikibase search. Although data analytics is

important for collecting evidence and making decisions [50,51], neither machine learning nor data are value-free [52]. Deciding which data items will be collected is a process that embeds a worldview (Weltanschauung). For example, is the information that is collected and searched about “abnormality” [53] or “participation restriction” [54]? In addition, the production of information is done for a certain purpose and hence is not value-free; for example, were certain data collected to understand the efficiency of people with disabilities in doing a task, or to determine if people with disabilities face discrimination in schools, universities, or the workplace? In this way, our virtual community is informed and designed in accordance with the human rights principles of the CRPD. Finally, the computer algorithms are value-loaded; indeed, predictive machine learning algorithms “learn” (ie, create models) from past data to predict events in the future based on current data. However, past data embeds values and biases from the past; for example, in 2018, Amazon’s AI-based recruiting tool was found to be sexist as it consistently suggested that male applicants be recruited [55], as it learned from resumes previously submitted to the company, which is in the male-dominated tech industry. Our choice to use the CRPD to build our ontology represents a first step in creating a search algorithm for disability data to be used by and for disability advocates and communities.

The current project requires many improvements and feature enhancements. First, the document data is semantified using a few relations, such as “has paragraph,” “has topic,” and “mentioned in.” Future work will include other relations, such as “comply with,” “supports,” and “has legislation” and update

the machine learning classifier to extract such semantic information from the uploaded document. Uploaded documents could be linked to their corresponding country as some documents are country- and region-specific. The current version of the application contains such properties in the interface; however, it is not embedded in the Wikibase data model. In the next phase, it is important to semantically link documents with countries and regions by importing countries and regions from Wikidata to the disability Wikibase.

Finally, the web application should be updated to include message broker architecture that allows server-side load balancing to handle large volumes of search requests and optimize the server’s performance.

Conclusions

The platform ontology was based on the CRPD to create a platform for equity-focused disability data that can be used by and for disability advocates and communities; the data includes advocacy reports and disability policy and legislation from specific jurisdictions or international contexts. The virtual community accepts 4 types of users: information producers, information consumers, validators, and administrators. Semiautomatic tagging and intelligent semantic search using natural language are enabled using AI techniques. The search engine supports bilingual (English and French) search and further enhancements are planned.

Multidisciplinary research in the domain of disabilities is challenging. An iterative process completed in a collaborative atmosphere allows team members to elucidate ambiguities and helps the team face challenges and provide solutions.

Acknowledgments

The authors acknowledge and thank The QA Company Inc for providing access to their intelligent search technology. The work reported in this paper was funded by a Social Sciences and Humanities Research Council (SSHRC) Insight Development Grant (grant number 430-2020-00819).

Authors' Contributions

CEM is the primary investigator who received the funds for this project. CEM and RG co-designed the study. CEM, RG, PM, and FM contributed equally to this phase of the project. AC and BK worked on the model triples and tested the tagging process; DD and RT developed the platform and machine learning algorithms. All approved the final version and agreed to be accountable for all aspects of the submitted paper. It is the understanding of the university and researchers that the Project Intellectual Property belongs to CEM, RG, PM, and FM.

Conflicts of Interest

PM is co-owner of The QA Company, which develops the QAnswer technology. All other authors declare no conflicts of interest.

References

1. Disability Data Advocacy Working Group. International Disability Alliance. 2019. URL: <http://www.internationaldisabilityalliance.org/content/disability-data-advocacy-working-group> [accessed 2021-07-15]
2. Disability Data Portal. Leonard Cheshire. 2020. URL: <https://www.disabilitydataportal.com/> [accessed 2021-07-15]
3. Rights Library - DRPI TRAINING. Disability Rights Promotion International (DRPI). 2016 Jul 06. URL: <https://drpitraining.research.yorku.ca/rights-library/> [accessed 2021-07-15]
4. UN Convention on the Rights of Persons with Disabilities. New York: Department of Economic and Social Affairs Disability, United Nations; 2006 Dec 13. URL: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html> [accessed 2021-07-15]

5. Loeb M. Disability statistics: an integral but missing (and misunderstood) component of development work. *Nord J Hum Rights* 2013;31(3):306-324 [FREE Full text] [Medline: [26925181](#)]
6. Abualghaib O, Groce N, Simeu N, Carew MT, Mont D. Making Visible the Invisible: Why Disability-Disaggregated Data is Vital to “Leave No-One Behind”. *Sustainability* 2019 May 31;11(11):3091. [doi: [10.3390/su11113091](#)]
7. Rioux M, Parekh G. Disability, Rights Monitoring, and Social Change. Toronto: Canadian Scholars' Press; 2015.
8. World Report on Disability. Geneva, Switzerland: World Health Organization; 2011. URL: https://www.who.int/disabilities/world_report/2011/report.pdf [accessed 2021-10-28]
9. Casebolt MT. Availability and quality of global disability data: A commentary on the Demographic and Health Surveys. *Disabil Health J* 2021 Jan;14(1):100972. [doi: [10.1016/j.dhjo.2020.100972](#)] [Medline: [32792331](#)]
10. Reed NS, Meeks LM, Swenor BK. Disability and COVID-19: who counts depends on who is counted. *The Lancet Public Health* 2020 Aug;5(8):e423. [doi: [10.1016/s2468-2667\(20\)30161-4](#)]
11. UN General Assembly. Convention on the rights of persons with disabilities: Article 31. New York, NY: The United Nations; 2007. URL: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-31-statistics-and-data-collection.html> [accessed 2021-07-15]
12. Resources for Data Users. The Washington Group. 2021. URL: <https://www.washingtongroup-disability.com/resources/resources-for-data-users/> [accessed 2021-07-15]
13. Strengthening the collection of data on disability. World Health Organization. 2021. URL: <https://www.who.int/activities/collection-of-data-on-disability> [accessed 2021-10-13]
14. El Morr C. Introduction to Health Informatics: A Canadian Perspective. Toronto, Canada: Canadian Scholars' Press; 2018.
15. El Morr C, Ali-Hassan H. Analytics in Healthcare: A Practical Introduction. Cham, Switzerland: Springer International Publishing; 2019:105.
16. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One* 2019;14(2):e0212356 [FREE Full text] [doi: [10.1371/journal.pone.0212356](#)] [Medline: [30779785](#)]
17. Lake IR, Colón-González FJ, Barker GC, Morbey RA, Smith GE, Elliot AJ. Machine learning to refine decision making within a syndromic surveillance service. *BMC Public Health* 2019 May 14;19(1):559 [FREE Full text] [doi: [10.1186/s12889-019-6916-9](#)] [Medline: [31088446](#)]
18. Sahoo A, Pradhan C, Das H. Performance Evaluation of Different Machine Learning Methods and Deep-Learning Based Convolutional Neural Network for Health Decision Making. In: Rout M, Rout J, Das H, editors. *Nature Inspired Computing for Data Science*. Cham, Switzerland: Springer International Publishing; 2020:201-212.
19. Chowdhary K. Natural Language Processing. In: *Fundamentals of Artificial Intelligence*. New Delhi, India: Springer India; 2020:603-649.
20. Edgcomb JB, Zima B. Machine Learning, Natural Language Processing, and the Electronic Health Record: Innovations in Mental Health Services Research. *Psychiatr Serv* 2019 Apr 01;70(4):346-349. [doi: [10.1176/appi.ps.201800401](#)] [Medline: [30784377](#)]
21. Velupillai S, Epstein S, Bittar A, Stephenson T, Dutta R, Downs J. Identifying Suicidal Adolescents from Mental Health Records Using Natural Language Processing. *Stud Health Technol Inform* 2019 Aug 21;264:413-417. [doi: [10.3233/SHTI190254](#)] [Medline: [31437956](#)]
22. Al-Natsheh HT, Martinet L, Muhlenbach F, Rico F, Zighed DA. Semantic Search-by-Examples for Scientific Topic Corpus Expansion in Digital Libraries. : IEEE; 2017 Presented at: IEEE International Conference on Data Mining Workshops (ICDMW); November 2017; New Orleans, LA, USA p. 747-756. [doi: [10.1109/icdmw.2017.103](#)]
23. Al-Natsheh H, Martinet L, Muhlenbach F, Zighed D. UdL at SemEval-2017 Task 1: Semantic Textual Similarity Estimation of English Sentence Pairs Using Regression Model over Pairwise Features. In: *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*.: The Association for Computational Linguistics; 2017 Presented at: The 11th International Workshop on Semantic Evaluation (SemEval-2017); August 3-4, 2017; Vancouver, Canada p. 115-119. [doi: [10.18653/v1/s17-2013](#)]
24. Bast H, Buchhold B, Haussmann E. Semantic Search on Text and Knowledge Bases. *FNT in Information Retrieval* 2016;10(1):119-271. [doi: [10.1561/15000000032](#)]
25. Gromann D. Neural language models for the multilingual, transcultural, and multimodal Semantic Web. *SW* 2020 Jan 31;11(1):29-39. [doi: [10.3233/sw-190373](#)]
26. Al-Natsheh HT. Text Mining Approaches for Semantic Similarity Exploration and Metadata Enrichment of Scientific Digital Libraries. Lyon, France: L'Université de Lyon; 2019.
27. Vrandečić D, Krötzsch M. Wikidata: a free collaborative knowledgebase. *Commun ACM* 2014 Sep 23;57(10):78-85. [doi: [10.1145/2629489](#)]
28. React. React: A JavaScript library for building user interfaces. 2021. URL: <https://reactjs.org/docs/getting-started.html> [accessed 2021-07-15]
29. MySQL Workbench: Enhanced Data Integration.: Oracle; 2021. URL: <https://www.mysql.com/products/workbench/> [accessed 2021-07-15]
30. MediaWiki.: Wikipedia Inc; 2020 May 26. URL: <https://www.mediawiki.org/wiki/MediaWiki> [accessed 2021-10-27]

31. Ehrlinger L, Wöß W. Towards a Definition of Knowledge Graphs. : CEUR-WS.org; 2016 Presented at: The 12th International Conference on Semantic Systems - SEMANTICS 2016; September 12-15; Leipzig, Germany. [doi: [10.1007/978-1-349-19066-9_2](https://doi.org/10.1007/978-1-349-19066-9_2)]
32. Accessing your Knowledge via Natural Language. QAnswer. 2021. URL: <https://www.qanswer.eu/> [accessed 2021-07-15]
33. Diefenbach D, Giménez-García J, Both A, Singh K, Maretditors. QAnswer KG: Designing a Portable Question Answering System over RDF Data. In: The Semantic Web. Cham, Switzerland: Springer International Publishing; 2020 Presented at: The Semantic Web, The 17th International Conference, ESWC 2020; May 31–June 4, 2020; Heraklion, Crete, Greece. [doi: [10.1007/978-3-030-49461-2_25](https://doi.org/10.1007/978-3-030-49461-2_25)]
34. Diefenbach D, Migliatti P, Qawasmeh O, Lully V, Singh K, Maret P. QAnswer: A Question Answering prototype bridging the gap between a considerable part of the LOD cloud and end-users. : Association for Computing Machinery; 2019 Presented at: The World Wide Web Conference; 13-17 May 2019; San Francisco, CA, USA. [doi: [10.1145/3308558.3314124](https://doi.org/10.1145/3308558.3314124)]
35. Python. 2021. URL: <https://www.python.org/> [accessed 2021-07-15]
36. Alshdaifat E. Hierarchical Ensemble Classification: Towards the Classification of Data Collections that Feature Large Numbers of Class Labels. Liverpool, UK: The University of Liverpool; 2015.
37. Joulin A, Grave E, Bojanowski P, Mikolov T. Bag of Tricks for Efficient Text Classification. : Association for Computational Linguistics; 2017 Presented at: The 15th Conference of the European Chapter of the Association for Computational Linguistics; April 3-7, 2017; Valencia, Spain. [doi: [10.18653/v1/e17-2068](https://doi.org/10.18653/v1/e17-2068)]
38. McKie JX. PyMuPDF (Version 1.18.19). 2021. URL: <https://github.com/pymupdf/PyMuPDF> [accessed 2021-10-15]
39. GitHub. 2021. URL: <https://github.com/> [accessed 2021-07-15]
40. Blei D, Ng A, Jordan M. Latent dirichlet allocation. *J Mach Learn Res* 2003;3:993-1022.
41. Zaveri A, Rula A, Maurino A, Pietrobon R, Lehmann J, Auer S. Quality assessment for Linked Data: A Survey. *SW* 2015 Mar 17;7(1):63-93. [doi: [10.3233/sw-150175](https://doi.org/10.3233/sw-150175)]
42. El Morr C, Maret P, Dinca-Panaitescu M, Rioux M, Subercaze J. Virtual community life cycle: A model to develop systems with fluid requirements. In: *Virtual Community Building and the Information Society: Current and Future Directions*. Hershey, PA, USA: IGI Global; 2012:1-14.
43. El Morr C, Maule C, Ashfaq I, Ritvo P, Ahmad F. A Student-Centered Mental Health Virtual Community Needs and Features: A Focus Group Study. *Stud Health Technol Inform* 2017;234:104-108. [Medline: [28186024](https://pubmed.ncbi.nlm.nih.gov/28186024/)]
44. El Morr C, AlHamzah M, Ng P, Purewal A, Al-Omran M. Knowledge of peripheral arterial disease: Results of an intervention to measure and improve PAD knowledge in Toronto. *Vascular* 2017 Oct 25;25(5):479-487. [doi: [10.1177/1708538116689355](https://doi.org/10.1177/1708538116689355)] [Medline: [28121280](https://pubmed.ncbi.nlm.nih.gov/28121280/)]
45. Gorman R, Maret P, Creighton A, Kundi B, Muhlenbach F, Buettgen A, et al. The Potential of an Artificial Intelligence for Disability Advocacy: The WikiDisability Project. *Stud Health Technol Inform* 2021 May 27;281:1025-1026. [doi: [10.3233/SHTI210338](https://doi.org/10.3233/SHTI210338)] [Medline: [34042833](https://pubmed.ncbi.nlm.nih.gov/34042833/)]
46. El Morr C, Eftychiou L. Evaluation Frameworks for Health Virtual Communities. In: *The Digitization of Healthcare*. London, UK: Palgrave Macmillan; 2017:99-118.
47. Elbattah M, Arnaud E, Gignon M, Dequen G. The Role of Text Analytics in Healthcare: A Review of Recent Developments and Applications. 2021 Presented at: The 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC); February 11-13, 2021; Vienna, Austria. [doi: [10.5220/0010414508250832](https://doi.org/10.5220/0010414508250832)]
48. Devlin J, Chang M, Lee K, Toutanova K. Pre-training of deep bidirectional transformers for language understanding. 2019 Presented at: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT); June 3-5, 2019; Minneapolis, MN, USA.
49. Beltagy I, Cohan A, Lo K. Pretrained contextualized embeddings for scientific text. Stroudsburg, PA, USA: The Association for Computational Linguistics; 2019 Presented at: The 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP); November 3-7, 2019; Hong Kong, China. [doi: [10.18653/v1/d19-1371](https://doi.org/10.18653/v1/d19-1371)]
50. Laudon K, Laudon J. *Management Information Systems: Managing the Digital Firm*. New York, NY: Pearson; 2017.
51. Pearlson K, Saunders C, Galletta D. *Managing and Using Information Systems A Strategic Approach*. Hoboken, NJ, USA: John Wiley & Sons; 2016.
52. Kelleher J, Namee B, D'Arcy A. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. Cambridge, Massachusetts, USA: MIT Press; 2015.
53. *The International Classification of Impairments, Disabilities, and Handicaps*. Geneva, Switzerland: World Health Organization; 1976. URL: <https://apps.who.int/iris/handle/10665/41003> [accessed 2021-10-28]
54. *Health topics - Disabilities*. Geneva, Switzerland: World Health Organization URL: <https://www.who.int/topics/disabilities/en/> [accessed 2021-07-15]
55. Dastin J. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. 2018. URL: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> [accessed 2020-01-16]

Abbreviations

AI: artificial intelligence
API: application programming interface
BERT: Bidirectional Encoder Representations from Transformers
CRPD: Convention on the Rights of Persons with Disabilities
DRPI: Disability Rights Promotion International
LDA: latent Dirichlet allocation
NLP: natural language processing
RDF: Resource Description Framework

Edited by G Eysenbach; submitted 03.09.21; peer-reviewed by C Slightam, JA Benítez-Andrades, M Elbattah; comments to author 27.09.21; revised version received 13.10.21; accepted 15.10.21; published 05.11.21.

Please cite as:

El Morr C, Maret P, Muhlenbach F, Dharmalingam D, Tadesse R, Creighton A, Kundi B, Buettgen A, Mgwigwi T, Dinca-Panaitescu S, Dua E, Gorman R

A Virtual Community for Disability Advocacy: Development of a Searchable Artificial Intelligence-Supported Platform
JMIR Form Res 2021;5(11):e33335

URL: <https://formative.jmir.org/2021/11/e33335>

doi: [10.2196/33335](https://doi.org/10.2196/33335)

PMID: [34738910](https://pubmed.ncbi.nlm.nih.gov/34738910/)

©Christo El Morr, Pierre Maret, Fabrice Muhlenbach, Dhayananth Dharmalingam, Rediet Tadesse, Alexandra Creighton, Bushra Kundi, Alexis Buettgen, Thumeka Mgwigwi, Serban Dinca-Panaitescu, Enakshi Dua, Rachel Gorman. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.11.2021. 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

Exploring the Barriers to and Motivators for Using Digital Mental Health Interventions Among Construction Personnel in Nigeria: Qualitative Study

Janet Mayowa Nwaogu^{1*}, MSc; Albert P C Chan^{1*}, PhD; John A Naslund^{2*}, PhD; Carol K H Hon^{3*}, PhD; Christopher Belonwu^{4*}, MSc; Jackie Yang^{1*}, PhD

¹Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong

²Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, United States

³School of Civil Engineering and Built Environment, Queensland University of Technology, Queensland, Australia

⁴Gribs Integrated Services Limited, Port Harcourt, Nigeria

* all authors contributed equally

Corresponding Author:

Janet Mayowa Nwaogu, MSc

Department of Building and Real Estate

The Hong Kong Polytechnic University

Kowloon

Hong Kong

Phone: 852 52243597

Email: janet.nwaogu@connect.polyu.hk

Abstract

Background: Work-related stress in the construction industry increases the prevalence of depression and anxiety among personnel. In low-resource settings such as Nigeria, construction personnel face high demands and severe working conditions but only have a few services to address their mental health needs. With emerging research showing that digital interventions can be used to self-manage mental health across diverse settings, there may be new opportunities to support construction personnel in the construction industry.

Objective: This study aims to determine the use of digital interventions for mental health management among construction personnel in Nigeria and to explore the factors that facilitate or impede the use of these interventions.

Methods: This qualitative study explored the perspectives of a convenience sample of 62 construction personnel. The data were subjected to inductive content analysis.

Results: A total of 6 barrier and 3 motivator themes were identified and categorized into 2 groups. The barrier themes were subcategorized into barriers to adoption and barriers to persistent use, whereas the motivator themes were subcategorized into intrinsic and extrinsic motivators. Lack of awareness and knowledge about the interventions may constitute a barrier to adoption and use. Participants frequently reported concerns regarding their effectiveness and usability.

Conclusions: This study provides an understanding of the design needs required to facilitate sustained self-management of mental health based on the experiences and expectations of construction personnel with digital interventions.

(*JMIR Form Res* 2021;5(11):e18969) doi:[10.2196/18969](https://doi.org/10.2196/18969)

KEYWORDS

mental health; construction personnel; digital technology; digital intervention; barriers; motivators; mobile phone

Introduction

Background

Construction personnel are reported to have a high incidence of common mental health problems [1-3], which results from

high levels of stress owing to the nature of their work and working conditions. These include poor remuneration [4], working in extreme temperatures [5], long working hours, job insecurity, role ambiguity [6], demanding jobs [7], work-life imbalance [8], and organizational injustice [9,10]. In the Netherlands, the prevalence rates of depression and

posttraumatic stress disorder symptoms among construction supervisors were 18% and 20%, respectively, whereas among skilled workers the rates were 11% and 7%, respectively [11]. In Nigeria, the prevalence rates of depression, anxiety, and suicidal ideation among construction supervisors were 55.1%, 14.8%, and 9.2%, respectively, whereas among tradesmen the rates were 74.5%, 36.4%, and 14.6%, respectively [12,13]. Generally, the weighted prevalence rates of depression, anxiety, and suicidal ideation in Nigeria were 5.5%, 3.5%, and 7.28%, respectively [14,15]. As untreated mental health problems pose a significant risk for suicidality [16] and other adverse health effects, there is an urgent need for mental health management among personnel in the construction industry and allied professions.

For lower middle-income countries such as Nigeria, there has been an increase in attention toward personnel's mental health and well-being needs, with recent efforts to mitigate stress reactions such as depression among construction personnel [4,10,17,18]. As the burden of mental health problems is predicted to increase over time [19], with an estimated 100 million mobile broadband users and 149 million active Global System for Mobile Communication subscribers in Nigeria, the country ranks highest in the internet market in Africa [20]; therefore, leveraging digital technology may yield more flexible, cost-efficient, and convenient methods to support the mental health of construction personnel. For instance, technology has proven to be potentially useful in reducing health management costs and challenges associated with mental health problems [21]. Through digital technologies such as smartphones, there may be opportunities for workers to manage stress responses and seek support for their mental health concerns when needed [1,22], while reducing possible barriers and stigma associated with accessing face-to-face services delivered on-site.

Workplace mental health interventions (eg, secondary and tertiary interventions) are used to mitigate and treat mental ill health [23] and overcome barriers to the use of psychological therapies delivered in mental health facilities. However, such workplace interventions are limited due to the need for in-person treatment, barriers to receiving and accessing care, stigma, shortage of mental health professionals, and high cost of service [1,24,25]. This is especially relevant among employees in the construction sector, as their roles are often mobile and involve excessive work hours [1]. Digital mental health interventions designed to meet specific needs, for example, managing stress, mood, anxiety, mental fitness, and well-being delivered through digital technologies such as smartphones, wearables, websites, or computers provide scalable interventions to improve health and health care [26].

Despite the surge in apps and websites, digital mental health interventions are underutilized [27]. A reason for the low adoption of interventions in mental health management in the workplace may be due to commercially available interventions that are not best suited to the workplace setting, user needs, and context [1]. In addition, these interventions often suffer from an abrupt exit by users without achieving the intended health goals [28]. To promote a sustained use of digital interventions until the desired outcome is achieved, identifying barriers and motivators for continued use is crucial. Such insights can inform

changes that may be needed for the design and efficacy of the intervention; without that, designers will largely rely on intuition [29]. This means identifying barriers to and motivators for the use of digital mental health interventions will impart valuable design information to support the development of efficient and user-friendly interventions [30].

Barriers and Motivators

Stiles-Shield et al [29] used card-sorting tasks to evaluate the barriers to using and engaging smartphone apps in managing depression among 20 potential users in the United States. The identified barriers to use were concerns about efficacy and functionality (ie, issues about effectiveness, misfit of feature to need, and crashing of the app), data privacy, insufficient feedback, and cost of data plans. Carolan et al [27] studied barriers to and facilitators of user engagement in a digital intervention aimed at psychoeducation and stress management among employees recruited from different sectors. Using semistructured interviews, the study deduced that the facilitators of digital intervention use and engagement included content, design, motivational emails (ie, quotes and advice), information on well-being, confidentiality, and support from the organization.

The use of digital interventions in construction industries, particularly wearable technologies for health and safety, has been studied [31,32]. However, these studies neither geared toward secondary intervention nor explored barriers to and motivators for use on an individual basis. Studies on occupational health have considered the barriers to and motivators for using digital mental health interventions among male-dominated professions in Australia [22,27]. Nonetheless, there remains a dearth of literature regarding high-risk populations, such as construction personnel in Nigeria. It is also unknown what constitutes the barriers and motivators to use such interventions among them. There is limited information on how construction personnel in low-resource countries use commercially available interventions to manage their mental health, what forms of barriers and motivators they experience, and whether their desired recovery goals are achieved.

Objectives

Examining users' experiences in specific contexts with digital mental health interventions could provide new information to better design future interventions. It is unknown how construction personnel in a low-income country such as Nigeria use existing digital mental health interventions to manage their health and what forms of barriers and motivators they encounter. Therefore, this study aims to fill the existing gap by exploring the satisfaction of construction personnel with publicly available digital interventions for mental health management in Nigeria. To achieve this aim, the following objectives were set: (1) to determine the barriers to use and engagement with digital mental health interventions among construction personnel in Nigeria and (2) to determine the motivators to use these interventions.

The ability to ensure engagement with the interventions after uptake will potentially allow construction personnel to cope with work stress and enhance positive mental health, especially if appropriate engagement theories and user needs ground the

development of such interventions. This can inform future design improvements, enabling acceptability and sustained use to achieve the intended mental fitness and health goals. This study adds to a limited body of knowledge by describing the factors that influence the probable uptake and use of digital mental health interventions by construction personnel in a lower middle-income setting. This study defines barriers as reasons or obstacles that prevent the use of digital interventions for mental health management. At the same time, motivators refer to catalysts for use, purpose, or factors that enhance the application or desire to use a digital mental health management intervention [33].

Methods

Theoretical Background

The Unified Theory of Acceptance and Use of Technology (UTAUT2) was chosen as the framework to guide this study as it is robust in incorporating previous acceptance models and is considered a highly sensitive model for explaining consumers' variance in the use of technology [34,35]. The UTAUT2 was adopted in this study through a bottom-up inductive approach to aid the interpretation of results and guide the design of future digital interventions. The UTAUT2 incorporates 7 constructs (ie, performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit) and 4 moderators (ie, age, gender, experience, and voluntariness) [36]. The first 5 constructs of the UTAUT2 relate to the enjoyment derived from using a product [37].

Sample and Measurement

This study used an open-ended survey, a qualitative technique. This technique was applied as it is often used to solicit feedback on products and design needs [38,39]. In the digital intervention literature, qualitative techniques are widely used as part of formative research efforts aimed at gaining insights from users about intervention design, content, and engagement [38]. Similarly, the qualitative technique was used to elicit feedback on digital interventions for mental health management from construction industry end users.

This technique is suitable as it provides a broader picture of the subject and provides an avenue for participants to express their concerns and feedback about the topic [40]. Such feedback can inform future intervention designs for the industry, whereby the industry can leverage such technologies to maintain mental fitness and promote good mental health among its workforce. Construction personnel engaged in supervisory and monitoring roles in Nigeria formed the target population for this survey and were determined based on a convenience and snowball sampling method.

The respondents included site supervisors (ie, site engineers) and project managers based on a log collected from the Nigerian Institute of Building. These groups of construction personnel were considered for 3 reasons: (1) involvement in building production management; (2) ease of eliciting the required information; and (3), should the needs of the construction personnel be met, leveraging digital interventions for universal prevention and secondary mental health interventions across

the entire construction workforce would be easier. The research methodology involved 2 steps: (1) survey questions were drafted based on literature reviews and (2) a pilot study was conducted with industry practitioners. The Mobile Application Rating Scale, a validated and reliable scale used for assessing the quality of health apps [41], was also used in this study. As this study explored the use of digital mental health interventions and not only apps, these questions were adapted for use in this study.

The survey questions were written in English Language and elicited demographic characteristics such as age, gender, and years of experience. Other information obtained was related to the type of digital mental health intervention that participants may have used, length of use, the intended goal of the intervention, and barriers to and motivators for the use of digital interventions. The questionnaire also contained a statement regarding informed consent and data confidentiality. As this survey pertained to adoption, satisfaction, and dissatisfaction with digital interventions used for mental health management, detailed information about the profession of respondents was not required. Before administering the questionnaire, the survey questions were piloted and reviewed by a professor with >15 years' experience in various aspects of occupational health and safety and a public health professional with expertise in health promotion and mental health. The Human Subjects Ethics Sub-Committee of Hong Kong Polytechnic University approved this study.

Piloting was performed to conduct face and content validity, assess readability, clarify language, and establish the appropriateness of the developed content within the intended context of use [42]. Suggestions made included soliciting information on gender, years of experience, and social media use, whereas some questions were reworded. The final survey is presented in [Multimedia Appendix 1](#) [27,41,43,44]. The survey questions were then developed using Google Forms, and the link was sent via an email to the respondents using Google Docs and the Mail Merge add-on by Quicklution.

Data Collection

The questionnaire was disseminated to the construction personnel via their organizational or personal email addresses. The Mail Merge add-on for Google Docs facilitated sending bulk and personalized emails to the respondents. An email introducing the study and containing the link to the survey was sent to the professionals. The first page of the survey link contained informed consent and confidentiality statements. On the basis of convenience sampling, a total of 335 emails were sent to the construction personnel in Nigeria between December 2018 and July 2019, while they were also encouraged to send out the survey link to their colleagues engaged in on-site building production and management.

A total of 62 responses were received following the convenience and snowball sampling. By responding to the questionnaire, this implies that 62 respondents gave informed consent. However, the number of responses were comparable with studies by Carolan and de Visser [27] and Anderson et al [43] on digital health interventions based on 18 and 22 responses, respectively. Comparable response rates (10%-20%) are common for surveys

conducted using emails or web-based mediums [45]. Although it is difficult to ascertain the total number of surveys sent out and the corresponding response rate due to the snowball sampling, Bowen et al [46] recorded a response rate of 7% following a web-based survey on work stress, stress effect, and coping in the construction industry.

The low number of responses may also be attributed to the research area being mental health management and due to stigma, it is possible that potential respondents thought that digital mental health interventions are meant only for persons with mental ill health symptoms. Responses were obtained from the respondents using Google Forms. To protect the confidentiality of respondents, the survey questions did not elicit contact details and affiliation of respondents in the survey form, and all responses collected through Google Forms were kept anonymous.

Data Analysis

Inductive content and thematic analysis of the retrieved data were supported using MAXQDA version 18.1.1 software for mixed methods analysis developed by VERBI GmbH. To ensure the reliability of the data analysis, the themes and categories were drawn in 2 stages to avoid bias; the codes, themes, and categories were also checked and agreed upon by at least 3 of the authors. First, the context or interpretation of the data was later refined using themes in the existing literature. The word cloud, word frequency, phrase finder, and categorization of survey data features were used as appropriate for supporting content and thematic analysis.

The word cloud and frequency tools were used to identify the most used words; the font size of such words was bold compared with others [47]. After identifying the most used words, they were read in line with the sentences they appeared in and autocoded appropriately. This implies, for instance, when asked, “Do you use social media? If yes, kindly mention those you mostly use.” The responses included “yes, Facebook, WhatsApp, LinkedIn” (Respondent NG #1); “FB, read news and latest post from my friends” (Respondent NG #3). On performing the analysis, words such as “Facebook, WhatsApp, LinkedIn, FB, and, read, news” were identified. The words were then read in line with the sentences in which they appeared to determine their worth. For example, “and, read, news” were not autocoded,

whereas other words were autocoded. “FB” was also coded to “Facebook.”

To avoid bias, the phrase finder (word combination) was used to identify possible themes and autocoded for longer sentences. Finally, to check for mistakes and refine the coding, the *categorize survey data* was used, after which refinement was performed manually. All autocoded themes in each question were refined in line with themes from previous studies using the *categorize survey data* tool. Subsequently, all segments were edited appropriately following the themes identified from previous literature or using a new theme as relevant to the statements. Concluding on themes during autocoding or refinement was also done with reference to the sentence in which they appeared as well as themes identified in Stiles-shield et al [29] and Simblett et al [30], followed by the descriptive analysis. The codes before refinement are shown in [Multimedia Appendices 2 and 3](#).

Results

Overview

Of the 62 respondents, 56 (90%) were male and 6 (10%) were female, and the mean age was 40 (SD 13.18) years. All respondents owned a smartphone and were active social media users. When asked which social media platforms they mostly used, respondents indicated that they used Facebook, WhatsApp, Instagram, Twitter, LinkedIn, and Telegram. Of the 62 respondents, only 24 (39%) had ever used digital interventions for mental health management. This implies that less than half of the respondents used their mobile devices or computers for issues related to their mental health.

Information was sought on the type of digital mental health interventions used by the respondents ([Table 1](#)). It was deduced that the interventions used included mobile-based apps (14/24, 58%) and web-based apps (10/24, 42%). It was further revealed that the respondents used mental health interventions for stress, depression, anxiety management, and mental fitness. In addition to using digital mental health interventions, 26% (16/62) of the respondents reported using activity or fitness trackers and other wearable technology. The wearable technologies used include activity trackers and bracelet technology aimed at maintaining mental health and fitness through health and wellness interventions.

Table 1. Demographics and how construction personnel in Nigeria use digital interventions for their mental health (N=62).

Demographic information	Participants, n (%)
Work experience (years)	
1-5	18 (21)
6-10	18 (24)
11-20	12 (21)
>20	14 (33)
Respondents who use electronic gadgets (mobile devices or computers) for mental health–related issues	
Yes	24 (39)
No	38 (61)
Respondents who use digital interventions for their mental health (n^a=24)	
Type of mental health intervention used	
Mobile-based app	14 (58)
Web-based app	10 (42)
Respondents who use wearables (n=16)	
Type of wearable devices	
Activity trackers	15 (91)
Bracelet technology	1 (9)
Wearable technology connection or synchronization (n=16)	
Smartphone connected	15 (91)
Standalone	1 (9)
Aim or purpose of using such digital intervention (n=17)	
Anxiety	6 (25)
Depression	5 (21)
Stress management	4 (17)
Mental fitness	2 (8)

^an: the number of responses.

The wearables were mostly phone connected or synchronized, although a few others, including the bracelet, were standalone wearables. The activity or fitness trackers include different brands of Fitbit, Apple Watch, Smartwatch, and Garmin mediums, used for lifestyle interventions, including physical activity, sleep tracking, heart rate, and blood pressure monitoring.

Barriers to Using Digital Interventions for Mental Health Management

When asked, “If you do not use your gadgets for anything concerning your mental health, would you like to?” Among the 38 participants who were yet to use such interventions, 28 (74%) expressed willingness to use their electronic gadgets for mental health management. This signaled that barriers to the adoption of digital mental health interventions among the participating

construction personnel in Nigeria could be related to a lack of awareness or limited knowledge about digital mental health interventions. In addition, of the 28 respondents willing to adopt digital mental health interventions, 4 (14%) indicated needing them for depression and anxiety management.

Respondents mentioned dislikes about the interventions and reasons to stop use, highlighting barriers to the continual use of digital mental health interventions (Table 2). Of the 24 respondents who had used their gadgets for something concerning their mental health, 15 (63%) reported no dislikes that will cause them to stop using the interventions, whereas 9 (37%) expressed some dislikes and reasons for stopping use. These dislikes and reasons for discontinuing digital mental health interventions were grouped into 5 categories: usability, efficiency and effectiveness, high cost, boring due to lack of human interface, and time-consuming.

Table 2. Linking the findings to the Unified Theory of Acceptance and Use of Technology (UTAUT2) framework.

Construct, themes	Responses	Identified UTAUT2 construct
Barriers		
Barrier to adoption		
Lack of awareness and knowledge	<ul style="list-style-type: none"> • ___^a 	<ul style="list-style-type: none"> • N/A^b
Barriers to persistent use		
Usability	<ul style="list-style-type: none"> • Navigation problems due to having too many options • Quite complicated • Technical malfunction • Battery life • Synchronization demands 	<ul style="list-style-type: none"> • Effort expectancy, facilitating conditions
Efficiency and effectiveness	<ul style="list-style-type: none"> • Not logical or practical • Not correct • Not accurate or not efficient • To obtain a better one 	<ul style="list-style-type: none"> • Performance expectancy
High cost	<ul style="list-style-type: none"> • Subscription cost • Data usage 	<ul style="list-style-type: none"> • Price value
Boring due to lack of human interface	<ul style="list-style-type: none"> • Robotic in nature and boring 	<ul style="list-style-type: none"> • Facilitating conditions
Lack of time	<ul style="list-style-type: none"> • Time consuming to input answers to all questions asked 	<ul style="list-style-type: none"> • Effort expectancy
Motivators		
Intrinsic and extrinsic		
Efficiency and effectiveness	<ul style="list-style-type: none"> • It is OK (ie, good) • It gives correct answers • Good and better mental health (ie, improved health) • Serve the purposes for which it is meant for • Give accurate results when monitoring health status • Stress reduction <ul style="list-style-type: none"> • relaxes palpitations • reduces stress • Quick access to health monitoring <ul style="list-style-type: none"> • quick access to solution • provides fast access to monitor mental ill health symptoms 	<ul style="list-style-type: none"> • Performance expectation • Performance expectancy • Performance expectancy, social influence
Usability	<ul style="list-style-type: none"> • Easy to use • App working well 	<ul style="list-style-type: none"> • Facilitating conditions, effort expectancy
Extrinsic		
Providing motivations	<ul style="list-style-type: none"> • Client's prompt motivational pull sessions 	<ul style="list-style-type: none"> • Hedonic motivation

^aNot available.

^bN/A: not applicable.

Feedback from respondents on the barriers to the persistent use of digital mental health interventions includes the following:

I use an app for mood tracking, quite good, but the options make it quite complicated. The wearable device I use can be annoying due to of the synchronization time. I need to delete initial sync to the app sometimes and start all over, and it can be stressful. [Respondent #2]

The applications have too many options, navigation becomes too tedious. [Respondent #22]

It is sometimes not correct. The beginning question on checking depression seems like prepared to classify me into having depression always. Sometimes, I input the wrong answers, but it still says I have depression. Sometimes, I struggle with downtimes, but not always. [Respondent # 41]

The app works but some of the activities to relieve stress don't seem to be logical or practical to help the situation. [Respondent #13]

I use a paid app for my stress and anxiety, while it seems good for me and needs dedication, subscribing is a bit expensive because it is international, maybe a local content will be a bit cheaper. [Respondent #34]

It is so time-consuming. It can consume time as someone has to answer several questions. [Respondent #53]

Respondents noted they would stop using the interventions for the following reasons:

It was consuming my data, so I stopped using it. [Respondent #8]

It is expensive to maintain premium subscription on the app I use. [Respondents #34]

I want to get a more accurate one. I used a free one that is artificial intelligence empowered; it asked questions about stress and how my sleep was. I answered by typing how unrested I feel when I wake, and it said I understand, followed by, that is, a lot of questions. The reply was off point. It put me off. I stopped using immediately. [Respondent #41]

It is robotic, nobody to respond or communicate. It can be boring. It is not exact. [Respondent NG #14]

Motivators for Using Digital Interventions for Mental Health Management

To determine the motivators for using digital mental health interventions, information was elicited on likes for interventions and services or things that would boost continual use of the interventions (Table 1). Furthermore, additional features that users of mobile-based apps would like to see in future mental health apps were determined. Motivators are measures that would support customer satisfaction and engagement. These features could also facilitate the continued use of mobile-based applications and reduce premature dropout from digital interventions when used in mental health management.

When asked what they liked about the interventions, participants' responses were grouped into 2 categories (Table 2): *efficiency and effectiveness* and *usability*. Efficiency and effectiveness were also related to the ability of the interventions to aid stress reduction and quick access to health monitoring. Specifically, respondents indicated the following:

It is quite easy to use, though the questions are not straightforward. [Respondent #1]

Feedback from instructor. [Respondent #12]

Gives correct answers. [Respondent #17]

I used it to monitor my anxiety level. [Respondent #16]

It makes me quickly check up on what level I am at my mental health space, fast access to checking the level of depression symptoms, though sometimes not completely accurate. [Respondent #41]

Quick access to solution. [Respondent #22]

It reduces stress. [Respondent #21]

It relaxes me from palpitations [Respondent #27]

Respondents were asked what services or features would encourage them to continually use the interventions until the desired health goal was achieved. The respondents indicated that ensuring the *usability, efficiency and effectiveness* of the interventions, and their ability to *provide motivational pulls* would boost sustained use. Respondents highlighted the following services or features to encourage engagement:

An app that works well. [Respondents #13]

Accurate results, it should be more accurate. [Respondents #18]

If it serves the purpose to which it is meant, then it is worth continuing. [Respondents #22]

Client's prompt motivational pull sessions. [Respondents #12]

When asked about additional features to incorporate into a digital mental health app, respondents indicated music, educative posts, motivational posts (ie, quotes), news alerts, comedy skits, and other health features, for example, some respondents mentioned the following:

Educative posts on health and some on work. [Respondent #1]

Meditational music. [Respondent #45]

Music. [Respondent #53]

I want an app that will be more than just mental health, an Artificial Intelligence-based app that is educative on other stress-related issues. [Respondent #58]

Discussion

Principal Findings

This study shows that key barriers to using digital mental health interventions among construction personnel in Nigeria include efficiency and effectiveness, usability, high cost, boring due to lack of human interface, and lack of time. On the basis of these results, lack of awareness of and knowledge about digital mental health interventions, if not attended to, may constitute a barrier to adoption. Hence, the barriers can be categorized as barriers to adoption and barriers to persistent use. Regarding the motivators to use, respondents mentioned usability, efficiency and effectiveness, and motivational features as important reasons they liked the interventions or features that could boost user engagement. It is evident that these are foundational characteristics that digital interventions need to possess to ensure user engagement and persistent use. Owing to the shortage of studies on this topic among lower middle-income economies and construction personnel, comparisons made in the discussion are drawn from evidence in the general body of knowledge, especially from populations in high-income countries.

Barriers to and Motivators for Using Digital Mental Health Interventions

Performance Expectancy

Performance expectancy measures were related to both barriers and motivators to using the interventions among construction personnel. They included efficiency and effectiveness measures, such as giving a better result, practicability to meet desired mental fitness and health needs, stress reduction ability, and quick access to health monitoring. In the case where the expectations were not met, it appeared a barrier to use, whereas if otherwise, it was a motivator for continued use. The ability of these interventions to enhance stress reduction and quick access to health monitoring can be described as an intrinsic motivator for using them. Generally, people use their digital devices to conduct a web search for information about their mental health, understand their experiences, and learn strategies to cope effectively [48]. These digital interventions allow for regular health assessments and quick access to resources for health-related information [49]. They can also be efficient and easy to use [50].

Meeting efficiency and effectiveness need is critical, as health care consumers are more likely to adopt and use electronic health technologies that can deliver the intended goals. In this study, some respondents had to stop using mental health apps due to efficiency and effectiveness concerns. There is a general concern over the ability of digital mental health interventions to meet intended treatment needs [29]. It is essential to conduct thorough quality testing for all features or services before deployment to facilitate engagement with digital mental health interventions. A crucial step to meeting effectiveness demand is for developers of future digital mental health interventions to adequately address the outlined concerns and needs that constitute barriers to use.

Bakker et al [51] suggested that such efficiency and effectiveness concerns could be linked to a lack of experimental validation for many publicly available digital mental health interventions. The ineffectiveness of the interventions may have resulted from a one-size-fits-all approach, which makes digital interventions less flexible and personalized [52]. Future digital interventions for the construction industry and allied professions in lower-middle-income settings such as Nigeria would benefit from ensuring that the gold standard for validation, which is a randomized controlled trial, is followed to ensure efficacy.

Facilitating Conditions and Effort Expectancy

Similarly, facilitating conditions and effort expectancy needs were both barriers and motivators to use. The facilitating conditions construct was related to usability and feelings that the interventions are boring due to lack of human interaction and personalized feedback, whereas concerns about usability and lack of time were related to effort expectancy. Usability appeared as a motivator and barrier. This finding is consistent with previous research, including a study that showed that boredom while using digital interventions to manage health negatively impacts user engagement [43]. To solve concerns about the lack of feedback, the integration of human support in interventions is a potentially valuable strategy [29] and may

assist in eliminating the robotic feeling and perception of ineffectiveness. Digital interventions for mental health delivered on the web (ie, web-based) that incorporate human guidance or coaching appear to yield better results and user engagement than unguided interventions [27,29]. A significant obstacle to the use of electronic health services by potential users is the absence of support or resources, which would allow them to use such services properly [35].

Usability concerns were related to a technical malfunction, which may have resulted from a sudden failure in the functionality of mobile apps. In this case, mental health apps were reported to fail during use, thus restricting the proper use of the interventions. Difficulties arising from connecting wearable devices with apps, the disappearance of apps (from the digital interface or dashboard), system freezing, and loss of power are significant factors that can affect the usability of digital interventions in managing health [30]. Another factor that affected the usability of digital mental health interventions was navigation demands, as some interventions were reported to have too many navigation options, making them complex to use.

Ensuring usability cannot be overlooked because construction personnel are more likely to use interventions that are easier to grasp when managing mental health [35]. To address usability concerns in mental health apps, it is important that apps are user tested and determined to be of good quality before deployment. A detailed guide for proper installation and use can also be provided to solve the functionality, synchronization, and navigation demands. It was also deduced that a large amount of time required to input information on feelings into digital mental health interventions was worrisome. This finding was consistent with that of Peng et al [44] in that the time required to input responses on feelings can be reduced by putting measures in place to satisfy the effort expectancy construct.

Price Value

This study reveals that the price value construct of UTAUT2 was related to high costs. High cost was identified as a barrier, as respondents mentioned that the cost of subscribing to mental health-related apps and internet access were deterrents for use. It included dislike for the cost involved in subscribing for the paid version of mental health apps as most were from a context outside Nigeria and involved elevated exchange costs and the cost for internet data. Furthermore, the cost of internet access (ie, data package) for deploying mental health apps [29] is a consistently reported barrier to using such interventions to manage mental health. Consumers of a product usually consider the cost related to acquiring a product or service; thus, price value is an essential determinant of use and acceptance [35]. Therefore, context-based interventions guided by strong psychological principles should be developed for a lower middle-income country such as Nigeria and its construction industry to address high cost concerns. Designers should ensure that efficiency and effectiveness are maintained while meeting affordability. This would mean that although extra (secondary) features can differ between subscription packages, the primary features necessary for effective mental health management should be the same and meet the required standard.

Most apps require a web-based connection, especially in cases of linked services, such as listening to meditation music. It would be necessary that designers include a download button, to enable downloading real-time services such as meditation music or e-coach session over a wireless network, which can be followed when a user is no longer connected to wireless access. This is especially important in low-resource settings such as Nigeria, where wireless connectivity can be highly variable. Designers should address these barriers in future digital interventions designed for the Nigerian context, as this would help ensure sustained engagement with digital interventions toward reaching the desired health goal and the promotion of improved mental health.

Hedonic Motivation

The need for quick motivational pull-ups and feedback could also serve as important extrinsic motivators and help solve the negative impact of social influence. For example, taking web-based interventions delivered through social media groups. Although such forums have helped in mental health management, concerns about the quality of information and materials accessed through such platforms have been highlighted [48]. The role of motivating users of such interventions is critical and should not be taken for granted. Integrating quick motivational pulls and feedback can increase their effectiveness and use. For online support groups, moderators should quickly attend to the concerns or issues raised. This will also help ensure efficiency and effectiveness, offering more significant potential to boost adoption and engagement.

Desired Additional Features for Mental Health Apps

Respondents indicated that comedy skits, music, educational posts, motivational posts, and news were additional features that they would like to see incorporated into future mobile-based apps. There is evidence showing that to increase employees' engagement with digital mental health interventions, motivational quotes are often sought [27]. Carolan et al [27] reported that half of the participants in their study affirmed receiving motivational quotes and well-being information, preferably every Monday morning. Such motivational messages are indicated to improve mental health, as people who choose to receive them show significant symptom reduction [53].

In the case where such interventions are personalized for delivery in an industrial setting, educational (informative) posts relating to occupation, health, and well-being can potentially improve engagement. The desire for comedy skits was a recurrent feature desired among the respondents and beams light in the context of Nigeria. The respondents reiterated that this feature, along with its effectiveness, would facilitate continual use. Wilkins et al [54] noted that humor and laughter impart health and well-being. In the Nigerian context, comedy skits are a popular medium of humor, which generates laughter that helps mitigate depression and other stress reactions [55]. Agwu and Tiemo [54] reported that laughter and biofeedback are ways to cope with stress among construction personnel. Comedy is a big business in Nigeria, a menu served at nightclubs in Nigeria and has moved from stand-up to web-based skits [55,56]. This may be because they provide a way to distract Nigerians from the worsening socioeconomic conditions that stare them in the

face. Incorporating animated comedy skits related to stress, mental health, and well-being as an additional feature in future mental health apps can provide hedonic motivation, thereby facilitating persistent use while delivering its primary health goals.

Designers of digital interventions will need to use their discretion to determine whether it is practicable to incorporate these additional features into an intervention while targeting the primary outcomes of mental health management. Mobile app interventions for mental health promotion of construction personnel can be enhanced with motivational posts (ie, quotes), educational posts, comedy skits, and preloaded relaxation music. Additional user feedback should be obtained, and the effectiveness of these features should be validated before deployment.

Limitations

Despite the overall sample size being small, it was valuable for an exploratory study of this nature. It is important to note that these results act as a guide to inform future studies focused on developing and delivering digital mental health interventions in the construction industry in a lower-middle-income country. Efforts were made to solicit as many responses as possible. However, because it was a web-based survey without compensation for respondents, many individuals may not have been self-motivated to respond. Another reason might be that the study is related to mental health, and individuals may have declined to participate due to stigma issues or misconceptions about mental health or cultural beliefs prevalent in Nigeria. The reduced response rate could also be due to the perception of no previous or intended need for such digital interventions for mental health management.

Another limitation of this study was the use of an anonymous web-based survey and open-ended questionnaire, which did not allow a follow-up to probe into some responses. However, this methodology was adopted to ensure anonymity and explore diverse views among users within the Nigerian construction industry. Owing to the limited sample size, it was not possible to determine the influence of age and gender on digital intervention use and acceptability, as detailed by UTAUT2. Nevertheless, because this study was exploratory and not conclusive, the influence of the variables was not a necessary factor. Finally, owing to the size of this study, the results may not be generalizable to other low-resource countries, similar occupational groups, or construction men. Therefore, these results should be interpreted with caution. However, this study spurs research reasoning in this direction.

Conclusions

This study explored the use of digital mental health interventions for self-management among construction personnel in Nigeria. Several barriers and motivators to the use of digital interventions in mental health management were deduced from this study and linked to the UTAUT2 framework following the inductive content analysis process. The barriers comprised usability, efficiency and effectiveness, high cost, boring due to lack of human interface, and lack of time. The motivators were efficiency and effectiveness, usability, and motivation.

This study shows that efforts are needed to provide awareness of the use of digital interventions for mental health management among construction personnel. This awareness should emphasize the different outcomes related to such interventions, including stress management, adequate sleep management, and mental fitness, which would help maintain good health. The construction industry needs to leverage digital interventions to ensure the mental fitness of its workers. For instance, cognitive behavioral therapy is also a preventive technique and can be delivered through digital mental health interventions, thus helping to prevent mental ill health. Therefore, the adoption of digital mental health interventions will be a potentially important avenue for delivering mental health services and universal prevention interventions throughout the construction industry in low-resource settings.

Although most of the findings are consistent with previous findings on digital health interventions, this study has extended the body of knowledge on digital interventions for mental health management by exploring construction personnel in Nigeria. The study reported that low awareness and knowledge about digital mental health interventions may be a viable barrier to use, with stress reduction (ie, the ability to reduce stress) as an intrinsic motivator to use. Furthermore, salient features desired by study respondents as additions to mental health apps included comedy skits, music, educative posts (resources), and motivational posts (ie, quotes). The respondents did not want their apps to be just a health tool; they also desired it as an educational and motivational tool. Future research should examine the suitability of such educational resources, such as routine tips on health, safety, and well-being. The provision of motivational quotes would be more impactful if they are enhanced with personalized features and tailored to the needs

of unique users. Tailoring has proved to be highly effective in supporting health communication and education efforts [44].

Future research should examine how organizations can adequately influence the use of digital mental health interventions in the Nigerian construction industry without fear of stigmatization. This study has revealed that feedback from users of digital mental health interventions can inform design information. This exploratory study offers a first step toward identifying some of the barriers to and motivators for the use of digital interventions in lower middle-income countries such as Nigeria. The next step would be to conduct a quantitative survey based on the findings of this study, which will then allow for comparative analysis and more conclusive results to inform the field of research. There is a need to further examine the level of awareness and knowledge about digital mental health interventions among construction personnel in Nigeria. This will help determine how uptake is affected by the knowledge about the different mental health fitness abilities that such interventions hold.

The findings from this study suggest that attending to these barriers and motivators will assist in eliminating abrupt exits from digital interventions for mental health management. The ability to ensure engagement with the interventions will allow construction personnel to cope with work stress and enhance positive mental health, especially if appropriate engagement theories and user needs ground the development of such interventions. The results of this study will assist in the design of new digital mental health interventions to deliver to personnel in the construction industry and improve existing interventions. Therefore, there is a need to leverage digital mental health interventions within the Nigerian construction industry for mental health promotion.

Acknowledgments

The Hong Kong Polytechnic University financially facilitated this research as part of a PhD study. Hence, studies with related backgrounds but differ in scopes and methodologies may be produced.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Complete survey question.

[\[DOCX File, 22 KB - formative_v5i11e18969_app1.docx\]](#)

Multimedia Appendix 2

Initial themes on motivators based on the likelihood of using digital interventions in mental health management.

[\[DOCX File, 44 KB - formative_v5i11e18969_app2.docx\]](#)

Multimedia Appendix 3

Initial themes based on responses to barriers to using digital interventions in mental health management.

[\[DOCX File, 44 KB - formative_v5i11e18969_app3.docx\]](#)

References

1. Deady M, Johnston DA, Glozier N, Milne D, Choi I, Mackinnon A, et al. A smartphone application for treating depressive symptoms: study protocol for a randomised controlled trial. *BMC Psychiatry* 2018 Jun 1;18(1):166 [FREE Full text] [doi: [10.1186/S12888-018-1752-5](https://doi.org/10.1186/S12888-018-1752-5)]
2. Sunindijo R, Kamardeen I. Work stress is a threat to gender diversity in the construction industry. *J Constr Eng Manage* 2017 Oct 01;143(10):04017073. [doi: [10.1061/\(ASCE\)CO.1943-7862.0001387](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001387)] [Medline: [29859060](https://pubmed.ncbi.nlm.nih.gov/29859060/)]
3. Roche AM, Pidd K, Fischer JA, Lee N, Scarfe A, Kostadinov V. Men, work, and mental health: a systematic review of depression in male-dominated industries and occupations. *Saf Health Work* 2016 Dec 01;7(4):268-283 [FREE Full text] [doi: [10.1016/j.shaw.2016.04.005](https://doi.org/10.1016/j.shaw.2016.04.005)] [Medline: [27924229](https://pubmed.ncbi.nlm.nih.gov/27924229/)]
4. Oladinrin TO, Adeniyi O, Udi MO. Analysis of stress management among professionals in the Nigerian construction industry. *Int J Multidiscipl Curr Res* 2014;2:22-33 [FREE Full text]
5. Ibem EO, Anosike MN, Azuh DE, Mosaku TO. Work stress among professionals in building construction industry in Nigeria. *J Construct Econom Build* 2011 Sep 21;11(3):45-57 [FREE Full text] [doi: [10.5130/ajceb.v11i3.2134](https://doi.org/10.5130/ajceb.v11i3.2134)]
6. Yip B, Rowlinson S. Job burnout among construction engineers working within consulting and contracting organizations. *J Manage Eng* 2009 Jul 01;25(3):122-130. [doi: [10.1061/\(ASCE\)0742-597X\(2009\)25:3\(122\)](https://doi.org/10.1061/(ASCE)0742-597X(2009)25:3(122))]
7. Leung M, Liang Q, Olomolaiye P. Impact of job stressors and stress on the safety behavior and accidents of construction workers. *J Manage Eng* 2016 Jan 01;32(1):04015019. [doi: [10.1061/\(ASCE\)ME.1943-5479.0000373](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000373)]
8. Lingard H, Francis V, Turner M. Work time demands, work time control and supervisor support in the Australian construction industry. *Eng Const Arch Manag* 2012 Nov 09;19(6):647-665. [doi: [10.1108/09699981211277559](https://doi.org/10.1108/09699981211277559)]
9. Chih Y, Kiazad K, Cheng D, Capezio A, Restubog S. Does organizational justice matter? Implications for construction workers' organizational commitment. *J Manage Eng* 2017 Mar 01;33(2):04016043. [doi: [10.1061/\(ASCE\)ME.1943-5479.0000490](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000490)]
10. Bowen P, Edwards P, Lingard H. Workplace stress experienced by construction professionals in South Africa. *J Constr Eng Manage* 2013 Apr 01;139(4):393-403. [doi: [10.1061/\(ASCE\)CO.1943-7862.0000625](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000625)]
11. Boschman JS, van der Molen HF, Sluiter JK, Frings-Dresen MH. Psychosocial work environment and mental health among construction workers. *Appl Ergon* 2013 Sep 01;44(5):748-755. [doi: [10.1016/j.apergo.2013.01.004](https://doi.org/10.1016/j.apergo.2013.01.004)] [Medline: [23380530](https://pubmed.ncbi.nlm.nih.gov/23380530/)]
12. Nwaogu JM, Chan AP, Tetteh MO. Staff resilience and coping behavior as protective factors for mental health among construction tradesmen. *J Eng Design Technol* 2021 Feb 04: Ahead-of-print. [doi: [10.1108/JEDT-11-2020-0464](https://doi.org/10.1108/JEDT-11-2020-0464)]
13. Nwaogu JM. An integrated approach to improve mental health among construction personnel in Nigeria. Thesis and Dissertations - The Hong Kong Polytechnic University, Hong Kong. 2021. URL: <https://www.polyu.edu.hk/en/> [accessed 2021-09-30]
14. Adewuya AO, Atilola O, Ola BA, Coker OA, Zachariah MP, Olugbile O, et al. Current prevalence, comorbidity and associated factors for symptoms of depression and generalised anxiety in the Lagos State Mental Health Survey (LSMHS), Nigeria. *Compr Psychiatry* 2018 Feb 01;81(1):60-65. [doi: [10.1016/j.comppsy.2017.11.010](https://doi.org/10.1016/j.comppsy.2017.11.010)] [Medline: [29268153](https://pubmed.ncbi.nlm.nih.gov/29268153/)]
15. Adewuya AO, Ola BA, Coker OA, Atilola O, Zachariah MP, Olugbile O, et al. Prevalence and associated factors for suicidal ideation in the Lagos State Mental Health Survey, Nigeria. *BJPsych Open* 2016 Nov 02;2(6):385-389 [FREE Full text] [doi: [10.1192/bjpo.bp.116.004333](https://doi.org/10.1192/bjpo.bp.116.004333)] [Medline: [27990294](https://pubmed.ncbi.nlm.nih.gov/27990294/)]
16. Li H, Luo X, Ke X, Dai Q, Zheng W, Zhang C, et al. Major depressive disorder and suicide risk among adult outpatients at several general hospitals in a Chinese Han population. *PLoS One* 2017 Jun 01;12(10):e0186143 [FREE Full text] [doi: [10.1371/journal.pone.0186143](https://doi.org/10.1371/journal.pone.0186143)] [Medline: [29016669](https://pubmed.ncbi.nlm.nih.gov/29016669/)]
17. Ojo GK, Adeyeye GM, Opaole A, Kajimo-Shakantu K. Gender differences in workplace stress response strategies of quantity surveyors in Southwestern Nigeria. *Int J Build Pathol Adapt* 2019 Oct 14;37(5):718-732. [doi: [10.1108/IJBPA-10-2018-0084](https://doi.org/10.1108/IJBPA-10-2018-0084)]
18. Sorensen G, Nagler EM, Pawar P, Gupta PC, Pednekar MS, Wagner GR. Lost in translation: the challenge of adapting integrated approaches for worker health and safety for low- and middle-income countries. *PLoS One* 2017 Jun 1;12(8):e0182607 [FREE Full text] [doi: [10.1371/journal.pone.0182607](https://doi.org/10.1371/journal.pone.0182607)] [Medline: [28837688](https://pubmed.ncbi.nlm.nih.gov/28837688/)]
19. Charles CN, Msagati T, Swai H, Chacha M. Microalgae: an alternative natural source of bioavailable omega-3 DHA for promotion of mental health in East Africa. *Scientific African* 2019 Nov 1;6(1):e00187. [doi: [10.1016/j.sciaf.2019.e00187](https://doi.org/10.1016/j.sciaf.2019.e00187)]
20. Russon MA. How internet access is improving in Nigeria. *BBC News*. 2020. URL: <https://www.bbc.com/news/business-51377955> [accessed 2021-04-30]
21. Mohr DC, Duffecy J, Jin L, Ludman EJ, Lewis A, Begale M, et al. Multimodal e-mental health treatment for depression: a feasibility trial. *J Med Internet Res* 2010 Dec 19;12(5):e48 [FREE Full text] [doi: [10.2196/jmir.1370](https://doi.org/10.2196/jmir.1370)] [Medline: [21169164](https://pubmed.ncbi.nlm.nih.gov/21169164/)]
22. Deady M, Peters D, Lang H, Calvo R, Glozier N, Christensen H, et al. Designing smartphone mental health applications for emergency service workers. *Occup Med (Lond)* 2017 Aug 01;67(6):425-428. [doi: [10.1093/occmed/kqx056](https://doi.org/10.1093/occmed/kqx056)] [Medline: [28535246](https://pubmed.ncbi.nlm.nih.gov/28535246/)]
23. Joyce S, Modini M, Christensen H, Mykletun A, Bryant R, Mitchell PB, et al. Workplace interventions for common mental disorders: a systematic meta-review. *Psychol Med* 2016 Mar 01;46(4):683-697. [doi: [10.1017/S0033291715002408](https://doi.org/10.1017/S0033291715002408)] [Medline: [26620157](https://pubmed.ncbi.nlm.nih.gov/26620157/)]

24. Dickerson FB, McNary SW, Brown CH, Kreyenbuhl J, Goldberg RW, Dixon LB. Somatic healthcare utilization among adults with serious mental illness who are receiving community psychiatric services. *Med Care* 2003 Apr;41(4):560-570 [FREE Full text] [doi: [10.1097/01.MLR.0000053440.18761.F0](https://doi.org/10.1097/01.MLR.0000053440.18761.F0)] [Medline: [12665719](https://pubmed.ncbi.nlm.nih.gov/12665719/)]
25. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR Mhealth Uhealth* 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: [10.2196/12106](https://doi.org/10.2196/12106)] [Medline: [30470676](https://pubmed.ncbi.nlm.nih.gov/30470676/)]
26. 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 01;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/)]
27. 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/)]
28. Nahum-Shani I, Hekler EB, Spruijt-Metz D. Building health behavior models to guide the development of just-in-time adaptive interventions: a pragmatic framework. *Health Psychol* 2015 Dec 01;34S(1):1209-1219 [FREE Full text] [doi: [10.1037/hea0000306](https://doi.org/10.1037/hea0000306)] [Medline: [26651462](https://pubmed.ncbi.nlm.nih.gov/26651462/)]
29. Stiles-Shields C, Montague E, Lattie EG, Kwasny MJ, Mohr DC. What might get in the way: barriers to the use of apps for depression. *Digit Health* 2017 Jun 01;3(1):2055207617713827 [FREE Full text] [doi: [10.1177/2055207617713827](https://doi.org/10.1177/2055207617713827)] [Medline: [29942605](https://pubmed.ncbi.nlm.nih.gov/29942605/)]
30. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. *J Med Internet Res* 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: [10.2196/10480](https://doi.org/10.2196/10480)] [Medline: [30001997](https://pubmed.ncbi.nlm.nih.gov/30001997/)]
31. Aryal A, Ghahramani A, Becerik-Gerber B. Monitoring fatigue in construction workers using physiological measurements. *Auto Construct* 2017 Oct 01;82(1):154-165. [doi: [10.1016/j.autcon.2017.03.003](https://doi.org/10.1016/j.autcon.2017.03.003)]
32. Hwang S, Lee S. Wristband-type wearable health devices to measure construction workers' physical demands. *Auto Construct* 2017 Nov 01;83(1):330-340. [doi: [10.1016/j.autcon.2017.06.003](https://doi.org/10.1016/j.autcon.2017.06.003)]
33. Passey ME, Longman JM, Robinson J, Wiggers J, Jones LL. Smoke-free homes: what are the barriers, motivators and enablers? A qualitative systematic review and thematic synthesis. *BMJ Open* 2016 Mar 17;6(3):e010260 [FREE Full text] [doi: [10.1136/bmjopen-2015-010260](https://doi.org/10.1136/bmjopen-2015-010260)] [Medline: [26988351](https://pubmed.ncbi.nlm.nih.gov/26988351/)]
34. Heselmans A, Aertgeerts B, Donceel P, Geens S, Van de Velde S, Ramaekers D. Family physicians' perceptions and use of electronic clinical decision support during the first year of implementation. *J Med Syst* 2012 Dec 01;36(6):3677-3684. [doi: [10.1007/s10916-012-9841-3](https://doi.org/10.1007/s10916-012-9841-3)] [Medline: [22402980](https://pubmed.ncbi.nlm.nih.gov/22402980/)]
35. Tavares J, Oliveira T. Electronic health record portal adoption: a cross country analysis. *BMC Med Inform Decis Mak* 2017 Jul 05;17(1):97 [FREE Full text] [doi: [10.1186/s12911-017-0482-9](https://doi.org/10.1186/s12911-017-0482-9)] [Medline: [28679423](https://pubmed.ncbi.nlm.nih.gov/28679423/)]
36. Venkatesh V, Thong J, Xu X. Unified theory of acceptance and use of technology: a synthesis and the road ahead. *J Asso Inform Syst* 2016 May;17(5):328-376 [FREE Full text] [doi: [10.17705/1jais.00428](https://doi.org/10.17705/1jais.00428)]
37. Idoga PE, Toycan M, Nadiri H, Çelebi E. Assessing factors militating against the acceptance and successful implementation of a cloud based health center from the healthcare professionals' perspective: a survey of hospitals in Benue state, northcentral Nigeria. *BMC Med Inform Decis Mak* 2019 Feb 19;19(1):34 [FREE Full text] [doi: [10.1186/s12911-019-0751-x](https://doi.org/10.1186/s12911-019-0751-x)] [Medline: [30782166](https://pubmed.ncbi.nlm.nih.gov/30782166/)]
38. Singer E, Couper MP. Some methodological uses of responses to open questions and other verbatim comments in quantitative surveys. *MDA J Quant Methods Survey Methodol* 2017;11(2):115-134. [doi: [10.12758/MDA.2017.01](https://doi.org/10.12758/MDA.2017.01)]
39. Vitale DC, Armenakis AA, Feild HS. Integrating qualitative and quantitative methods for organizational diagnosis. *J Mixed Methods Res* 2008 Jan 01;2(1):87-105. [doi: [10.1177/1558689807309968](https://doi.org/10.1177/1558689807309968)]
40. Miller AL, Dumford AD. Open-ended survey questions: item nonresponse nightmare or qualitative data dream? *Surv Pract* 2014 Oct 01;7(5):1-11. [doi: [10.29115/SP-2014-0024](https://doi.org/10.29115/SP-2014-0024)]
41. 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/)]
42. Afolabi MO, Bojang K, D'Alessandro U, Ota MO, Imoukhuede EB, Ravinetto R, et al. Digitised audio questionnaire for assessment of informed consent comprehension in a low-literacy African research population: development and psychometric evaluation. *BMJ Open* 2014 Jun 24;4(6):e004817 [FREE Full text] [doi: [10.1136/bmjopen-2014-004817](https://doi.org/10.1136/bmjopen-2014-004817)] [Medline: [24961716](https://pubmed.ncbi.nlm.nih.gov/24961716/)]
43. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. *PLoS One* 2016 Jun 01;11(5):e0156164 [FREE Full text] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]
44. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. *BMC Public Health* 2016 Nov 14;16(1):1158 [FREE Full text] [doi: [10.1186/s12889-016-3808-0](https://doi.org/10.1186/s12889-016-3808-0)] [Medline: [27842533](https://pubmed.ncbi.nlm.nih.gov/27842533/)]
45. Fricker Jr RD. Sampling methods for web and e-mail surveys. In: *The SAGE Handbook of Online Research Methods*. Thousand Oaks, CA: SAGE Publications Ltd; 2008:195-217.

46. Bowen P, Govender R, Edwards P. Structural equation modeling of occupational stress in the construction industry. *J Constr Eng Manage* 2014 Sep 01;140(9):04014042. [doi: [10.1061/\(ASCE\)CO.1943-7862.0000877](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000877)]
47. Fielding J, Fielding N, Hughes G. Opening up open-ended survey data using qualitative software. *Qual Quant* 2012 May 5;47(6):3261-3276. [doi: [10.1007/s11135-012-9716-1](https://doi.org/10.1007/s11135-012-9716-1)]
48. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. *Psychol Psychother* 2019 Jun 01;92(2):277-297. [doi: [10.1111/papt.12222](https://doi.org/10.1111/papt.12222)] [Medline: [30924316](https://pubmed.ncbi.nlm.nih.gov/30924316/)]
49. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: a survey of youths' web-based resource preferences in Canada. *J Med Internet Res* 2014 Dec 17;16(12):293-295 [FREE Full text] [doi: [10.2196/jmir.3526](https://doi.org/10.2196/jmir.3526)] [Medline: [25519847](https://pubmed.ncbi.nlm.nih.gov/25519847/)]
50. Naslund JA, Aschbrenner KA, Scherer EA, McHugo GJ, Marsch LA, Bartels SJ. Wearable devices and mobile technologies for supporting behavioral weight loss among people with serious mental illness. *Psychiatry Res* 2016 Oct 30;244(1):139-144 [FREE Full text] [doi: [10.1016/j.psychres.2016.06.056](https://doi.org/10.1016/j.psychres.2016.06.056)] [Medline: [27479104](https://pubmed.ncbi.nlm.nih.gov/27479104/)]
51. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016 Mar 01;3(1):7-31 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
52. Scholten H, Granic I. Use of the principles of design thinking to address limitations of digital mental health interventions for youth: viewpoint. *J Med Internet Res* 2019 Jan 14;21(1):11528-11552 [FREE Full text] [doi: [10.2196/11528](https://doi.org/10.2196/11528)] [Medline: [31344671](https://pubmed.ncbi.nlm.nih.gov/31344671/)]
53. Whitton AE, Proudfoot J, Clarke J, Birch MR, Parker G, Manicavasagar V, et al. Breaking open the black box: isolating the most potent features of a web and mobile phone-based intervention for depression, anxiety, and stress. *JMIR Ment Health* 2015 Sep 20;2(1):e3 [FREE Full text] [doi: [10.2196/mental.3573](https://doi.org/10.2196/mental.3573)] [Medline: [26543909](https://pubmed.ncbi.nlm.nih.gov/26543909/)]
54. Wilkins J, Eisenbraun AJ. Humor theories and the physiological benefits of laughter. *Holist Nurs Pract* 2009;23(6):349-354 [FREE Full text] [doi: [10.1097/HNP.0b013e3181bf37ad](https://doi.org/10.1097/HNP.0b013e3181bf37ad)] [Medline: [19901609](https://pubmed.ncbi.nlm.nih.gov/19901609/)]
55. Adesoye RE. Phonological distortion as a humorous strategy in Folarin Falana's comedy skits. *Eur J Humour Res* 2018 Dec 30;6(4):60-74. [doi: [10.7592/EJHR2018.6.4.adesoye](https://doi.org/10.7592/EJHR2018.6.4.adesoye)]
56. Adetunji A. The interactional context of humor in Nigerian stand-up comedy. *Pragmatics* 2013 Mar 1;23(1):1-22. [doi: [10.1075/prag.23.1.01ade](https://doi.org/10.1075/prag.23.1.01ade)]

Abbreviations

UTAUT2: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 30.03.20; peer-reviewed by D Sugarman, B Schouten, J Kim; comments to author 10.07.20; revised version received 04.05.21; accepted 08.08.21; published 09.11.21.

Please cite as:

Nwaogu JM, Chan APC, Naslund JA, Hon CKH, Belonwu C, Yang J

Exploring the Barriers to and Motivators for Using Digital Mental Health Interventions Among Construction Personnel in Nigeria: Qualitative Study

JMIR Form Res 2021;5(11):e18969

URL: <https://formative.jmir.org/2021/11/e18969>

doi: [10.2196/18969](https://doi.org/10.2196/18969)

PMID: [34751652](https://pubmed.ncbi.nlm.nih.gov/34751652/)

©Janet Mayowa Nwaogu, Albert P C Chan, John A Naslund, Carol K H Hon, Christopher Belonwu, Jackie Yang. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 09.11.2021. 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 Fully Collaborative, Noteless Electronic Medical Record Designed to Minimize Information Chaos: Software Design and Feasibility Study

Jackson Steinkamp¹, MD; Abhinav Sharma², MD; Wasif Bala³, MD; Jacob J Kantrowitz⁴, MD, PhD

¹Hospital of the University of Pennsylvania, Philadelphia, PA, United States

²Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada

³Emory University Hospital, Atlanta, GA, United States

⁴Carney Hospital, Dorchester, MA, United States

Corresponding Author:

Jackson Steinkamp, MD

Hospital of the University of Pennsylvania

3400 Spruce Street

Philadelphia, PA

United States

Phone: 1 215 662 4000

Email: jacksonsteinkamp@gmail.com

Abstract

Background: Clinicians spend large amounts of their workday using electronic medical records (EMRs). Poorly designed documentation systems contribute to the proliferation of out-of-date information, increased time spent on medical records, clinician burnout, and medical errors. Beyond software interfaces, examining the underlying paradigms and organizational structures for clinical information may provide insights into ways to improve documentation systems. In particular, our attachment to the *note* as the major organizational unit for storing unstructured medical data may be a cause of many of the problems with modern clinical documentation. Notes, as currently understood, systematically incentivize information duplication and information scattering, both within a single clinician's notes over time and across multiple clinicians' notes. Therefore, it is worthwhile to explore alternative paradigms for unstructured data organization.

Objective: The aim of this study is to demonstrate the feasibility of building an EMR that does not use notes as the core organizational unit for unstructured data and which is designed specifically to disincentivize information duplication and information scattering.

Methods: We used specific design principles to minimize the incentive for users to duplicate and scatter information. By default, the majority of a patient's medical history remains the same over time, so users should not have to redocument that information. Clinicians on different teams or services mostly share the same medical information, so all data should be collaboratively shared across teams and services (while still allowing for disagreement and nuance). In all cases where a clinician must state that information has remained the same, they should be able to *attest* to the information without redocumenting it. We designed and built a web-based EMR based on these design principles.

Results: We built a medical documentation system that does not use notes and instead treats the chart as a single, dynamically updating, and fully collaborative workspace. All information is organized by clinical topic or problem. Version history functionality is used to enable granular tracking of changes over time. Our system is highly customizable to individual workflows and enables each individual user to decide which data should be structured and which should be unstructured, enabling individuals to leverage the advantages of structured templating and clinical decision support as desired without requiring programming knowledge. The system is designed to facilitate real-time, fully collaborative documentation and communication among multiple clinicians.

Conclusions: We demonstrated the feasibility of building a non-note-based, fully collaborative EMR system. Our attachment to the *note* as the only possible atomic unit of unstructured medical data should be reevaluated, and alternative models should be considered.

(JMIR Form Res 2021;5(11):e23789) doi:[10.2196/23789](https://doi.org/10.2196/23789)

KEYWORDS

electronic medical records; clinical notes; information chaos; information overload; clinician burnout; software design; problem-oriented medical record; medical records; electronic records; documentation; clinical; software

Introduction

Inefficient Electronic Medical Records: Why It Matters

Clinicians spend much of their workday using electronic medical records (EMRs), both getting data *in*—entering new information about patients—and getting data *out*—searching for and summarizing past information. Some studies estimate that ambulatory physicians spend approximately 50% of their work time navigating, searching, and entering information into the EMR [1]. Therefore, optimizing clinical information management systems is crucial for clinician efficiency, satisfaction, and high-quality patient care. Poorly designed information management systems lead to the proliferation of out-of-date or incorrect information [2-4], increased time spent searching medical charts [1,5,6], medical errors [3,4], and clinician burnout [5,7-10], while limiting the effective use of EMR data for individual or population-level applications.

Given the scope and impact of the problem, we must critically evaluate the core conceptual assumptions that underpin our current practices. For instance, are *notes* the best organizational unit for unstructured clinical data? Should clinicians in different teams always document in separate locations? Which data should be recorded and stored in a *structured* fashion, and which in an *unstructured* fashion? Historically, these questions have been informed less by principled usability concerns and more by regulatory requirements, vendor incentives, and holdover intuitions from the era of paper records, which have fundamentally different capabilities and limitations than electronic systems. A robust field of EMR usability studies would enable empirical examination of our intuitions regarding the optimal management of clinical data using controlled studies [11-16].

What Creates Inefficient EMRs: Chaos and Complexity

Patient care is exceedingly complex, and the practice of medicine has increasingly recognized this fact [17-19]. Medicine has heightened its focus on holistic care, population health, and the behavioral and socioeconomic causes of medical diseases. However, the EMR systems that have emerged in the past 20 years are inappropriately suited to clinicians' mental models and the realities of modern medical practice [20].

Within the EMR, clinical data are stored in multiple formats: unstructured clinical notes, semistructured reports, and structured lists of laboratory values, medications, problems, and allergies. Beasley et al [2] defined a framework, information chaos, for analyzing the information problems that arise in modern electronic charts and plague physicians. A physician's ability to diagnose and treat heavily relies on the interpretation of information, much of which is stored within the EMR, such as the subjective history, physical examination, vital signs, laboratories, and imaging findings. Information chaos is a theory that highlights the spectrum of difficulties incurred in processing

this large volume of information, which hinder timely access to this information. In addition to producing delays in care, these obstacles result in a substantial but nonessential expenditure of cognitive effort on the part of clinicians, ultimately contributing to higher rates of clinician burnout and medical error. The theory delineates five major hazards of information chaos: information overload, information underload, information scatter, conflicting information, and erroneous information. These hazards reinforce each other in overt and subtle ways (see [Multimedia Appendix 1](#) for additional details). A user-friendly EMR should incentivize behaviors that minimize information chaos to facilitate quick and accurate clinical decisions and summarization. In contrast, poorly designed systems and paradigms *increase* the likelihood of behaviors that lead to redundant, conflicting, scattered, and erroneous information, inhibiting excellent care.

Although EMR system designs exacerbate information chaos, a clearly delineated educational framework for teaching and assessing physician-learner EMR skills could ameliorate some of these issues. The reporter, interpreter, manager, and educator (RIME) framework provides an educational framework that could be readily extended to EMR competencies, as outlined by Stephens et al [21] (RIME/EMR scheme). They argue that the Accreditation Council for Graduate Medical Education core competencies are applicable to EMR skill development and clinical practice and that the importance of information technology in facilitating lifelong physician learning necessitates an examination and careful planning of educational strategies to develop effective EMR users. As physicians develop from reporters to educators clinically, their abilities to find relevant data, order appropriate tests, and document thoroughly and accurately within the electronic systems must improve to match. However, the lack of intelligent organization of the data, unnecessarily complicated ordering systems, and paper-era documentation tools impede the advancement of these skills. A user-friendly EMR, designed through the lens of the RIME/EMR scheme, could provide easier reporting tools; guide rails for learning to interpret, integrate, and plan based on clinical data; and decision support based on the best, up-to-date evidence in nearly real time. Such a system is reminiscent of Weed's [22] proposed system that *guides and teaches*.

What Creates Inefficient EMRs: Dealing With Structured and Unstructured Data

In modern EMR systems, limited classes of structured data are recorded and viewed within *non-note* interfaces. The historical distinction between which data are structured and unstructured is largely arbitrary and a historical consequence of the "Meaningful use" initiative and other regulatory initiatives. Storing information in a structured form facilitates data-specific interfaces (eg, medication reconciliation interfaces), clinical decision support (CDS) tools, and automated population-level analyses. However, entering data in a structured format is more onerous for the clinician than the large, unstructured, free-text

blocks such as those found in notes, particularly if the interface requires interruption of cognitive flow or multiple clicks to navigate between structured data entry points. Furthermore, individual clinicians are unable to define their own structured data elements to improve their workflows. We should question the traditional distinctions between structured and unstructured data. Specifically, the principle of designing to minimize information chaos can be used to guide a more principled understanding of structured data—historical precedent alone should not be the guide for system design.

How the Current Idea of the Note Contributes to EMR Inefficiencies

Bundles, defined by Gorman [23] as “organized, highly selective collections of information,” are used extensively by clinicians to manage information in health care settings characterized by uncertainty, frequent interruptions, and grave outcomes. In these settings, time and attention are limited, and interdisciplinary care is essential. Effective design and selection of information bundles is critical to effectively communicate, filter, and act on medical information.

Currently, the most common EMR paradigm for documenting unstructured information uses bundles called *notes*. A note is simply a bundle of text that contains a group of related but distinct clinical observations and decisions. Notes generally organize unstructured information in three major ways: (1) by time slice (each note focuses on events and reasoning from a discrete, limited span of time), (2) by clinical thread (each note focuses on particular clinical topics and omits others), and (3) by responsibility (each note’s writer is responsible for the entirety of the note’s content). These 3 principles of unstructured data organization are sensible (Table 1); however, a note, as currently conceived, is a poor choice of information bundle that directly creates information chaos.

Note-based organization causes information chaos at two major levels: (1) at the level of the single clinician duplicating information in multiple notes over time and (2) at the level of multiple clinicians or teams duplicating information in independent clinical threads.

Table 1. The 3 major ways that notes organize unstructured information into discrete bundles.

	Organizational paradigm		
	Time slice	Thread	Responsibility
Description	<ul style="list-style-type: none"> Each note holds information from a particular time slice of a patient’s history. 	<ul style="list-style-type: none"> Different sets of notes focus on sets of particular clinical topics (a thread) and omit or deemphasize others. 	<ul style="list-style-type: none"> The writer of each note is responsible for and simultaneously attests to all information within it.
Examples	<ul style="list-style-type: none"> Each progress note in a set of hospital notes holds information about a particular 24-hour period; each outpatient visit holds information about the period between the previous visit and that visit. 	<ul style="list-style-type: none"> A thread of outpatient progress notes focuses on the management of chronic problems; a thread of nursing notes in a hospitalization focuses on nursing care issues; and a thread of cardiology consult notes focuses on the patient’s cardiology history, exam, diagnoses, and treatments. 	<ul style="list-style-type: none"> Attending physicians and trainees may write separate notes or note segments containing the same information.
Utility	<ul style="list-style-type: none"> Enables sorting and filtering information by time 	<ul style="list-style-type: none"> Enables limited filtering by subject matter (for consult services, social work notes, and nursing notes) Enables entirely independent documentation by clinicians in different threads 	<ul style="list-style-type: none"> Enables clear assignment of responsibility for a particular set of facts and reasoning.
Source of information overload or duplication	<ul style="list-style-type: none"> Most of a patient’s medical information remains the same from time t to time t+1; however, it is often redocumented or copy-pasted forward from a previous note. 	<ul style="list-style-type: none"> Most of a patient’s medical information is shared among different clinical teams or threads; however, it is redocumented or copy-pasted in multiple independent note threads. 	<ul style="list-style-type: none"> Information may be redocumented multiple times or by multiple clinicians rather than simply <i>attesting</i> to one’s agreement with a colleague.
Source of information scatter	<ul style="list-style-type: none"> Difficult to track the course of a chronic problem over time (it either requires navigation among many notes or accumulative charting, where each note continues to grow in size and hold the entire history of the patient) Difficult to document small single-fact updates without making an entirely new note 	<ul style="list-style-type: none"> Difficult to see where clinicians disagree on a particular issue (history element and examination finding), as it requires navigation among multiple notes 	<ul style="list-style-type: none"> N/A^a

^aN/A: not applicable.

The first problem may be caused by 2 rival conceptions about what a note as time slice is supposed to be. On the one hand, a note can be seen as a *bundle of new updates*—only holding the information that has *changed* since the previous note. This minimizes information duplication but increases information scatter; tracking a patient's problem over time requires navigating across many notes. On the other hand, a note can be seen as a complete *current state of the patient*. In this view, anyone reading only a single note has a complete picture of the relevant information at that time. This practice minimizes information scatter but leads to charts bloated with duplicate information, making it difficult for a later reader to identify what has changed from time t to $t+1$, and makes it onerous to expunge errors from the chart. This conception of the note forces us into a trade-off between information scatter and information overload. In addition, information is commonly duplicated across multiple note threads, although most of the patient's clinical data are the same. The information that does differ is usually unintuitively scattered in different places, making it difficult to reconcile conflicts, thereby limiting the potential for meaningful collaboration within the EMR. In this study, we aim to develop a novel EMR system that uses a fully collaborative, dynamic, and problem-oriented approach to minimize information chaos within electronic patient charts. We aim to explore the pros and cons of a non-note paradigm in an attempt to expand the dialog around EMR usability.

Methods

Overview

We sought to design a system that breaks down the traditional barriers of structured and unstructured data based on the notion that the current organizational principle of *notes* is a major contributor to information chaos within the EMR. In particular, we believe that the rigid organization of unstructured data by *time slice*, *thread*, and *responsibility* leads to information overload and information scatter, which in turn may lead to wasted time and effort, medical errors, user dissatisfaction, and clinician burnout. We hypothesize that it is feasible to build a modern EMR that does not use notes as the major organizational units for unstructured data and allows clinicians the flexibility they require to perform an unpredictable job.

Design Principle: Designing to Minimize Information Chaos

To prevent information overload and scatter, the core operating principle of a clinician using our EMR is as follows: redocumentation is to be fundamentally disincentivized at every aspect of the system's design. Taken to its logical extreme, in an ideal system, a user should find it *absurd* or *unthinkable* to redocument the same information multiple times within the same patient chart. Details will be discussed in the *Results* section, but the key principles are listed below.

First, within an individual clinician or team's *thread* of documentation over time, the default assumption is that information will remain the same over time; information that changes is the *exception*. This is a sensible assumption, given that a patient's accumulated past medical history will rarely be

changed once recorded, and many aspects of a patient's current medical state will remain the same between 2 consecutive clinical notes. This means that our system's interfaces should be designed around *small-scale changes of individual data elements* without requiring redocumentation of the remainder of the unchanged data—something not possible under a note framework, which usually treats the *note* as the atomic unit of writable and viewable information. In contrast, we identify *individual data elements* as the atomic unit of writable information rather than the *note*, which is, by definition, an aggregate of many data elements. This is the first intuition that leads us to reject the *note* as an information organization principle.

Second, clinicians on *different* teams or services (ie, those who would traditionally use different note *threads*) should not be required to maintain separate copies of the information that does not differ among teams or services. We would seek to reduce the number of distinct clinical threads and incentivize the shared maintenance of an entirely *collaborative* chart wherever possible. This has numerous benefits, including reducing information overload, information scatter, and error propagation while facilitating more meaningful collaboration within the EMR. This requires dissolving the sharp boundary between notes written by different teams in favor of a default collaborative workspace. Although this may initially worry some readers, we will explain in detail how we deal with the concerns around responsibility tracking in the *Results* section. We believe that the benefits significantly outweigh the costs.

Third, wherever it needs to be recorded that information *has not changed* since the last documentation event, the fundamental mode of doing so will be *attestation*, not *redocumentation*. A clinician can attest to the current state of the information they are responsible for without typing it out again (or even copy-pasting it). This saves time and effort while reducing information overload. This principle also requires us to reject the *note* as the atomic information unit in favor of something more like a dynamic *living workspace*.

Design Principle: Baked-In CDS

Another key design principle, which our system shares with the traditional problem-oriented medical record [22], is the notion that medical records should *guide and teach*—that is, facilitate learning and CDS at the problem level. For instance, if a clinician indicates that a patient has congestive heart failure (CHF), the system should present the clinician with a framework for how to think about that problem, including relevant data to be collected, active diagnostic possibilities, actions to be considered (including diagnostic tests, therapeutic orders, consults, and dispositional concerns), and links to external information resources. These frameworks should be informed by other information already captured in the chart—for instance, if a cardiac history was already captured in the context of a coronary artery disease workup, that information should be viewable by default from within the CHF problem, removing the need to search for that information (and, if unchanged, the need to *redocument* it).

In this way, CDS is *baked-in* at the point of diagnosis and patient conceptualization, not merely at the point of structured data

entry (eg, medication orders). A good documentation system serves as a *mental scratchpad*, which mirrors clinician cognition while providing suggestions and support to overtaxed human brains [24,25]. Such systems would not only help trainees learn how to think about particular problems but also facilitate the incorporation of the latest treatment guidelines and research studies by clinicians at all levels of training. Given the large volumes of information required for modern medical practice, such a system would likely serve an important *checklist* role, freeing physicians to focus on more uniquely human cognitive and emotional concerns. A proper system would allow for sensible defaults at the institution level but would *always* enable clinicians to build their own custom individual templates for their own practices. Textual templates in existing note-based EMRs enable some of these capabilities but do not have the kind of granularity or conditional logic necessary to achieve the full potential of such a system, largely because of the constraints of a note-based organization.

Then, a necessary component of a non-note-based system is an engine that enables clinicians to build their own granular problem-based CDS and workflows. We describe our implementation of such a system in the *Results* section.

Software

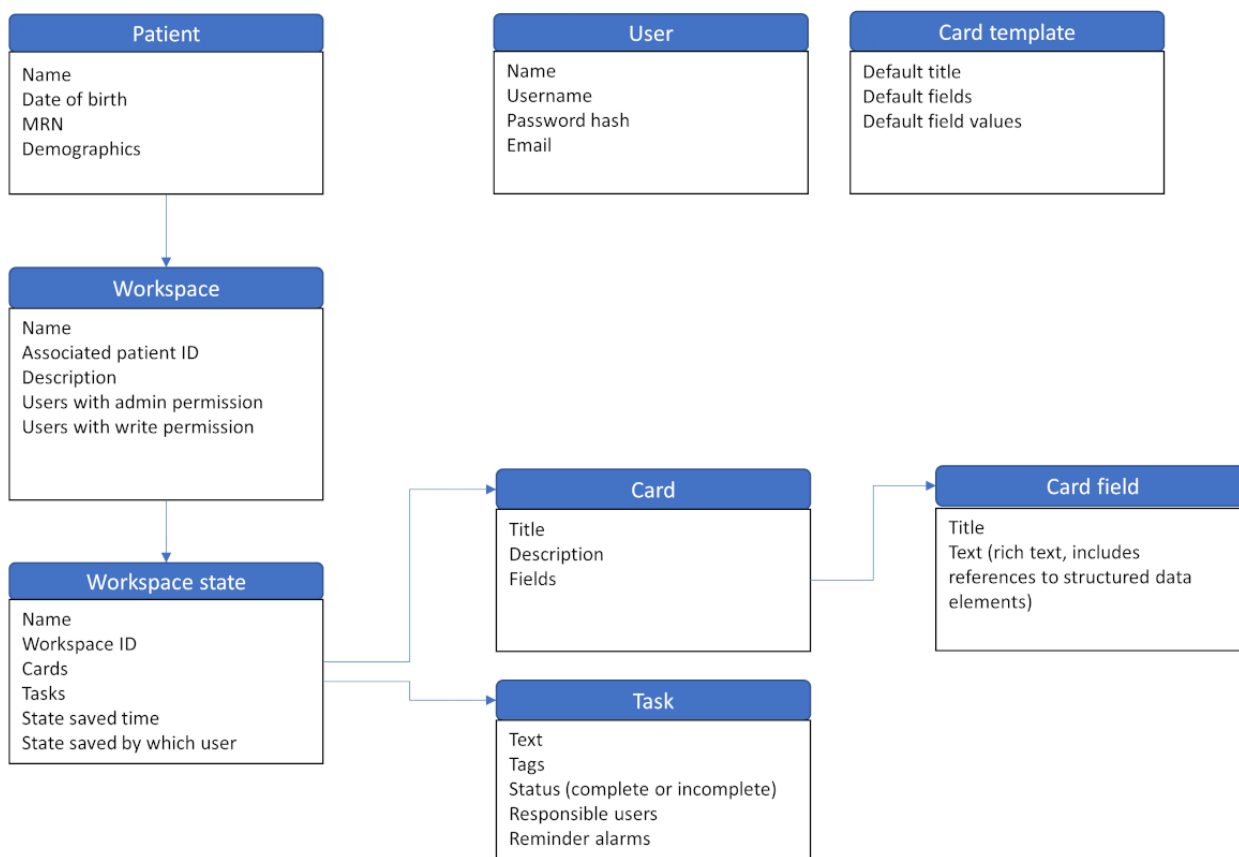
Our system is a web-based EMR system. All the server-side code was written using the Python language with the *flask* web framework. On the back end, we used the MongoDB database to store app data. The front end is written in JavaScript using the Vue framework. The SocketIO framework was used to facilitate real-time multiuser collaboration and push notification functionality. When an update to a patient’s chart is made, the server notifies all other active users of the change, enabling real-time collaboration. For further details on the implementation of real-time collaboration, see [Multimedia Appendix 2](#) [26]. The software was developed iteratively with multiple rounds of idea generation, user testing, and feature refinement.

Results

Overview

Below, we describe the major components of the software and their implementation. Each of these key features is designed to decrease information chaos—particularly the duplication and scatter of related information in disparate locations. [Figure 1](#) shows the overall relationships between the different conceptual entities and data structures in the MongoDB database.

Figure 1. Database diagram showing the organization of the major conceptual entities in the system. MRN: Medical Record Number.



Cards

Under a *note-based* paradigm, all information covering a particular time slice of a patient’s medical life (eg, a single day of hospitalization) must be stored together in one note, regardless of content. As discussed above, this organizational

scheme forces users to write notes that are overloaded with information or that are underloaded, with information scattered across many notes. Instead of the note as time slice approach, our system takes inspiration from work on the problem-oriented medical record [22] and organizes information by topic to facilitate charts that *teach and guide*. In our software, each topic

is represented by a *card*, which holds a set of information related to that topic. Each card consists of a title and multiple *fields* that provide additional structure within the topic (Figure 2). Cards are frequently used to capture individual medical problems; for instance, a CHF exacerbation card might hold fields representing the patient's history of present illness, relevant physical exam findings, outpatient clinicians, laboratory and imaging results, and treatment orders. However, cards are designed to be flexible enough to capture topics that do not neatly map to medical problems, such as "demographic data", "cardiology consult recommendations", "list of outpatient clinicians", "nursing comments", or "overnight updates" (Figure 3). This enables institutions, teams, or individual clinicians to develop cards customized to their own individual workflows. The fields are, by default, all free-text, retaining all the advantages of textual data while adding a small amount of flexible structure. Each card field's data are stored and updated separately in the database rather than every topic being stored together in a single block of free text. This enables granular viewing of individual cards or fields and their changes over time, in the case where a clinician wants to quickly see the history of a patient's CHF.

However, the default chart view shows all currently active cards for a patient regardless of how recently each piece of information has changed. This is sensible, as different medical information update at different frequencies. Even for the same problem, a patient's acute status may change daily, whereas their outpatient medications may change monthly, their

providers change still less often, and their demographics may never change. Therefore, clinicians can get a quick understanding of the current state of the patient without having to navigate back through numerous notes to find updates that occurred at different times. Our system shares these advantages with other forms of *problem-based* or *topic-based* charting.

By using cards as the basic organizational principle, we escape the trade-off between information overload and information scatter. Our system facilitates granular updates to individual topics without requiring large-scale redocumentation and information overload. For instance, if everything about a patient's CHF has remained the same from day 1 to day 2, the card does not need to be modified at all. Instead, the responsible clinician simply attests that no information has changed, which saves time and does not create a duplicate copy of the same data (as is the case with notes). Updates that consist of only 1 or 2 factual updates to the chart (ie, a telephone call to refill medication) require clinicians to only update a single card field. Unlike a note-based organization, which would scatter topically related information across multiple notes, the relevant data will be stored together regardless of when it was documented, enabling easier viewing of longitudinal changes. Therefore, the card-based system cuts down on both information overload and information scatter. This structure also enables readers (or automated systems) to quickly identify which cards and fields have changed from update to update and which have remained the same. We will discuss this further in the *Version History* section.

Figure 2. An example of a problem card template comprising a title, description, and list of free text data fields. Users build these problem card templates and instantiate them for a particular patient. The green bars represent structured data, which will automatically be pulled into the card when an instance of this card template is created. The cards add flexible, customizable structure to the topic-based documentation process.

Card Title:	Coronary Artery Disease (CAD)	🔒 +
Card Description:	Management and information for coronary artery disease.	
Field title:	Medications	🔍 🗑️
Field title:	Cholesterol	🔍 🗑️
	Lipid Panel	
Field title:	Percutaneous Interventions (PCI)	🔍 🗑️
Field title:	Catheterization Results	🔍 🗑️
Field title:	Cardiologist	🔍 🗑️
	Last Cardiologist	

Figure 3. An example of a card that does not map to a medical problem. The flexible structure of custom cards enables users to build their cards into the workflow. Note that ED course and PCP are structured data elements built by the user, which can be pulled automatically into multiple cards while retaining a single source of truth. ED: emergency department, HPI: history of present illness, PCP: primary care provider.

The cards also facilitate real-time collaboration. Multiple clinicians can work on individual cards or even different fields within the same card at the same time, reducing the incentive for individuals on the same team to create separate notes. Furthermore, card-based organization can reduce the number of separate clinical threads; ideally, inpatient consultants would be responsible for individual cards or fields within a general inpatient workspace, obviating the need for the consultant to redocument large amounts of the patient's history of present illness medication list, etc. Real-time collaboration is facilitated using web sockets, which allow for two-way communication between client and server; when a client makes any update to a card within a patient workspace, the server will push the update via the web socket, enabling other users to see the changes immediately. This facilitates the reconceptualization of the chart as a *dynamic living workspace*.

Structured Data

Structured data elements can be pulled into unstructured documentation using an @ operator inside any of the free text card fields, akin to a mention or hashtag system in other software—we call this the *structured data display operator* (Figure 4). For instance, if “@sodium” is typed inside a card field, a set of options related to the stored sodium laboratory values will appear. In its simplest form, a reference to the latest sodium value will be pulled into the text field (Figure 5) and displayed in a rich text format. Critically, this structured data value is not a disjoint piece of text but a reference that can be updated as new information arrives. If the user wants to track the latest value of sodium, and a new value of sodium is input

into the system, the textual reference will display a *refresh* icon, suggesting that the displayed value is no longer the latest value (Figure 6). The user can choose whether they wish to make use of the refresh icon to update their textual documentation. This is designed to prevent the propagation of incorrect values because a clinician forgot to update the text—a scenario that has the potential for significant medical error. We allow each structured data element to be displayed in multiple formats—a “last value” option, an “all values” option, a “specific value” option (which does not change with time), and a *graph* option, which enables a real-time updating graph of the latest values of a particular structured data element (Figure 7), such as a laboratory value. Note that these data elements are situated within the free text card fields, removing the requirement to navigate with multiple clicks to separate *graph viewing* interfaces and showing topic-relevant information in the same place—preventing information scatter, underload, and lost time.

To facilitate flexible, customizable workflows, we take the widest possible view of what counts as structured data by enabling individual users to define and create their own structured data elements, as well as enter any structured data value themselves (eg, laboratory values obtained from outside organizations). This includes traditional structured data elements, such as laboratory results, but is also designed to work with other numeric values (ejection fraction, pack-years of tobacco use, and risk scores) in addition to string values and categorical values (code status and the patient's current outpatient pulmonologist). This empowers individual clinicians to quickly improve their templating systems.

Figure 4. Use of the structured data display operator. Typing @ in a card field will create a menu for different structured data elements, which can be pulled into the text of a card field, including the ability to create graphs and lists. The user is free to define and customize their own structured data elements.

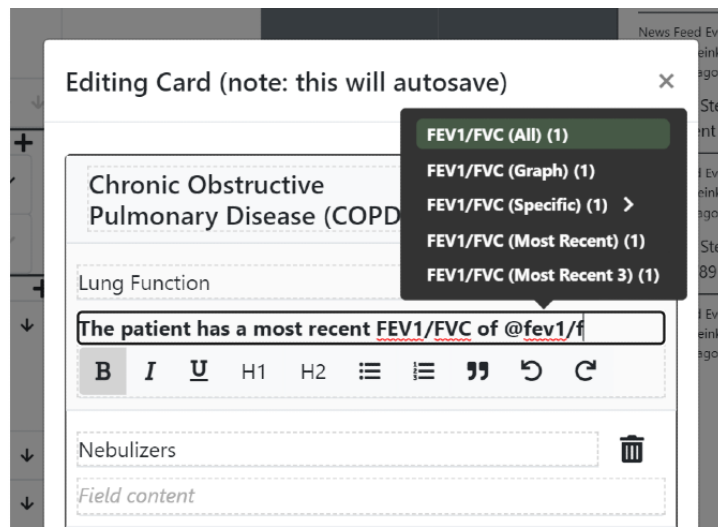


Figure 5. Clicking one of the menu options will pull data from the shared structured data pool. It is highlighted in green to indicate that it is an updatable structured data element, not a raw text.

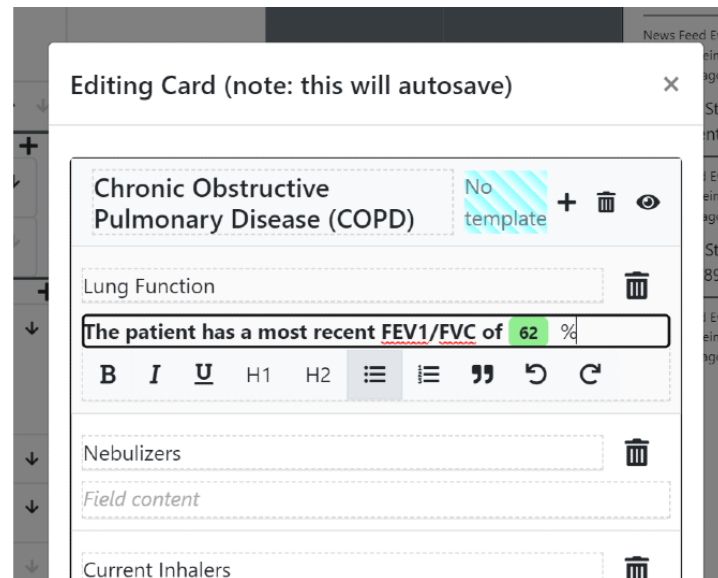


Figure 6. The refresh icon appears in a card if the underlying structured data element has changed. If the user wants to maintain a particular structured data value from a particular time (eg, the sodium on admission), the refresh icon will only appear if that value is updated. When a user wants to track the most recent values of a structured data field (eg, the most recent ejection fraction at the time of viewing that card), the refresh icon will appear if there are any more recent values for the same structured data element (eg, a more recent forced vital capacity than the one currently displayed).

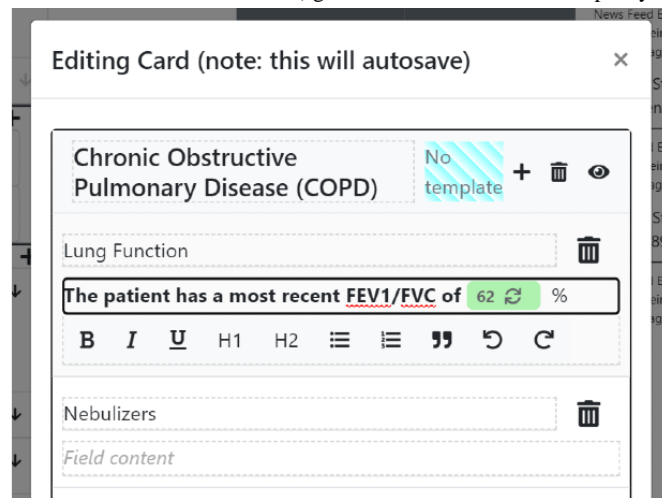
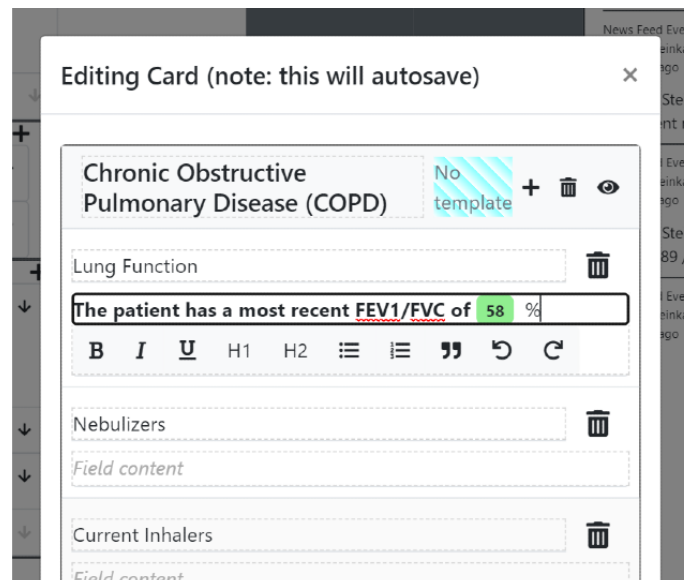


Figure 7. When the refresh icon is clicked, the data element in the card field updates to reflect the changes without requiring the user to view the structured data element.



Template Engine

Most patient encounters involve a small set of common medical problems, suggesting that a significant benefit is to be gained from templated workflows. In current EMR systems, *note templates* are used to realize this benefit. We implemented a much more granular templating system at the level of *card templates* (Figure 8). A card template consists of a set of prespecified fields with optional prespecified default text in each field. Inside a patient's workspace, cards can be quickly instantiated from the templates (Figure 9). This default text can include structured data elements, enabling the clinician to (for instance) autopopulate a reference to the latest ejection fraction when a card is instantiated from a template. This is another

advantage realized by a noteless EMR, and it is here where the potential of medical charts that *guide and teach* can begin to be realized. Trainees, in particular, may benefit from card templates, which provide a framework for thinking about a particular medical topic, including important diagnostic considerations, therapeutic decision points, and signs of patient status change, all of which can be included within a card template. Physicians play multiple roles, and these teaching medical charts can help us to be better clinicians, educators, and learners by providing perpetually up-to-date resources naturally within our existing workflows. Card templates can include embedded text from external resources such as guideline organizations and research studies, enabling not only clinical refreshers but also ongoing medical education.

Figure 8. Card templating engine, enabling the design of templated problem-specific fields, which can pull in structured data (part 1).

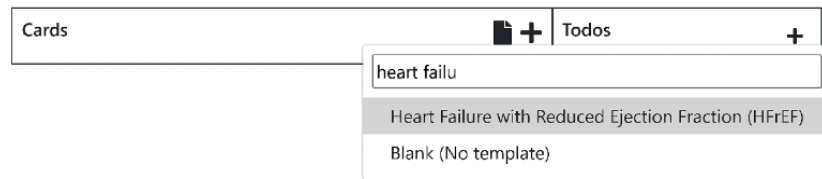
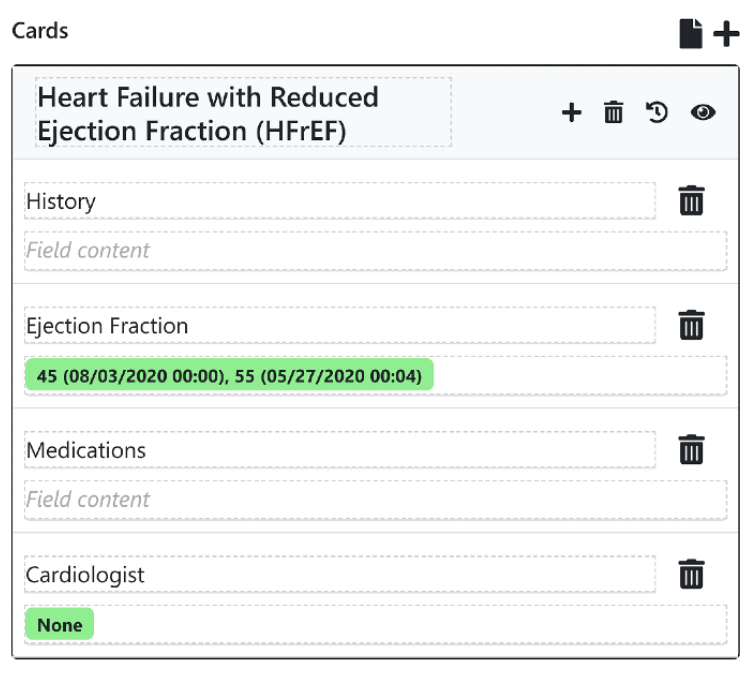


Figure 9. Card instantiated from template.



Such templates have the potential to make clinicians at all levels more efficient and prevent medical errors caused by forgetfulness or outdated information. In this way, CDS is *built into* the information organizational framework in a way that provides guidance more than it aims to correct. Traditional CDS is typically based on structured data, meaning the support is necessarily given after data are created, in reaction to a particular pattern of data (eg, medication interactions, allergies, and laboratory values). On the other hand, by instantiating a card template with a particular structure, the card instantiation itself becomes a form of CDS.

In addition to a simple menu-based selection system for instantiating card templates, our system includes a separate interface that enables fast autopopulation of cards from free text, such as a patient *one-liner*. A one-liner is a single-sentence patient summary that lists, among other things, a patient's relevant medical problems. Our system uses simple rules to identify which card templates should be created from a given block of text; however, more sophisticated natural language processing systems could be used to improve performance and handle more complicated linguistic phenomena. In the future, such a system could be used to import free text blocks directly from other systems that operate according to note-based organizational systems and quickly convert them to the more flexible card-based organization.

Workspaces

Although our system is designed to cut down on the number of distinct clinical threads and the information overload that results from multiple clinicians or teams redocumenting the same information, we understand that there are potential use cases where clinicians should work in separate *workspaces*. Therefore, we allow for the creation of separate workspaces for the same patient, which recaptures the notion of separate *threads*. The system requires only a single click to navigate from one workspace to another in the same patient's chart. Each workspace has specific user permissions: clinicians with write permission to a workspace can see the most up-to-date version of a workspace, even if it has not yet been attested. On the other hand, those with read permissions can only see the most recent *attested* version to prevent acting on incomplete information. For custom workflows, cards from one workspace can be *watched* from another. Watching a card creates a read-only reference to the original card in the new workspace and enables a clinician to view relevant information documented by someone else from another workspace. When the original card updates, the *watched* copy of the card will also update in real time. This is meant to replace the current behavior under a note-based system in which clinicians copy-paste information (eg, consult recommendations or a radiology report) from one text document into another to have all the information in one place. Although other design decisions of our system will ideally decrease the information scatter that motivates this behavior, card-watching

is nevertheless present to mitigate the creation of duplicate *unlinked* data copies.

Task Management

In many clinical care settings, relevant patient information is frequently passed between clinicians and staff through ancillary systems such as emails, pagers, SMS text messages, and separate handoffs. Our system includes a task management or *to-do list* tracking system as a core feature. Each workspace includes a dynamically editable, fully collaborative to-do list that enables task management for that patient. To-do lists can be assigned to particular users, can be associated with reminders accompanied by push notifications, and can be given custom user tags to enable filtering and searching functionality. These to-do lists can be viewed in aggregate across a group of patients, enabling efficient inpatient team and outpatient panel management.

Real-Time Collaboration

Notes are not an effective way to communicate when there is some urgency to the task; immediate messaging systems are more appropriate for this use case. Therefore, our system includes a *chat* functionality as the primary mode of communication. The chat system enables direct messaging among clinical users in the system and includes a group chat functionality, which is similar to the chat functionality present in other EMR systems. In addition, in our system, each patient workspace is associated with its own additional chat, so short notes or messages related to a particular patient can be stored. However, unlike many other chat systems, ours is public and, thus, auditable. The public nature of the chat enables any system user, including physician colleagues or learners, nurses, and other staff, to understand the clinical decisions that have been made for their patients and to learn from these conversations, as opposed to these happening in closed-door sessions.

Within the workspaces, we include a real-time cursor tracking feature: a user in a patient chart can see which other users are editing that workspace and where their cursor is located to prevent multiple users from editing the same information at the same time and facilitate real-time collaboration (eg, by serving as a jumping-off point for a chat conversation between a primary clinician and a consultant).

Although some may be hesitant to treat the chart as a collaborative workspace, based on fears of medicolegal reprisal or insufficient assignment of responsibility to individuals, we believe our system is capable of handling the responsibility assignment issue at least as well as note-based systems. Each edit (addition or deletion of a card, editing of a card title or field text, etc) is tracked in a separate logging table, enabling fine-grained assignment of responsibility for individual edits

to particular users. These logs also enable the ascription of responsibility for each piece of text in a workspace to particular users and particular time stamps. This feature can be helpful not only in cases of ascribing responsibility but also when a user wishes to orient themselves to a new chart and understand what information has been recently updated and what is in need of verification or updating.

Version History

Under our system, the default assumption is that most information *will not change* from update t to $t+1$. This intuition conceives of the chart as a dynamic living workspace and facilitates an intuitive version history system akin to that found in modern word-processing and other collaborative software. Each patient workspace consists of a set of states representing how the workspace has been updated over time by any number of clinicians working collaboratively. In particular, each workspace contains a single *current* state, which reflects the latest view of the workspace, and any number of past states, which reflects how the workspace looked at particular points in the past. Each state consists of a set of cards, to-do tasks, and structured data elements, as mentioned above. When a clinician wishes to *attest* to the state of a workspace, they click a *save* or *attest* button to indicate this. Similar to signing a note, attestation is designed to be used for final versions of documentation that will go into the medicolegal record. Attestation creates a locked past state that can no longer be edited by anyone but is added to the record of past states.

Users can scroll back and forth through previous workspace states to quickly see how information has changed over time (Figure 10). A prominent *show changes* button enables intuitive highlighting of cards, fields, and individual sentences (Figure 11), enabling quick focus on the small pieces of information that have changed from state t to state $t+1$. Although nothing in principle prevents a version history system from being implemented in note-based EMRs, our card-based organization maximizes its use by enabling individual cards or fields to be tracked over time. Software should be optimized for the most common use cases; tracking medical problems over time is perhaps the single most common use case of an EMR, but the *note* abstraction is ill-suited for it.

Clinicians can also export the current state of a workspace to a single block of text, enabling our system to be compatible and interoperable with note-based systems. For instance, clinicians could document in our system and export a workspace state as a note-like block of text to another documentation system. This would enable our system to be used as an adjunct to a note-based EMR and facilitate direct comparative studies of usability without requiring large-scale institutional changes.

Figure 10. Diagram of the overall organization of a patient interface for a fake patient. Each workspace enables separate clinical teams to keep separate threads on rare occasions when it is necessary. Each workspace stores a history of states; that is, the current state of the workspace and older states of the workspace (past versions). These can be scrolled through easily using the left and right arrows at the top.

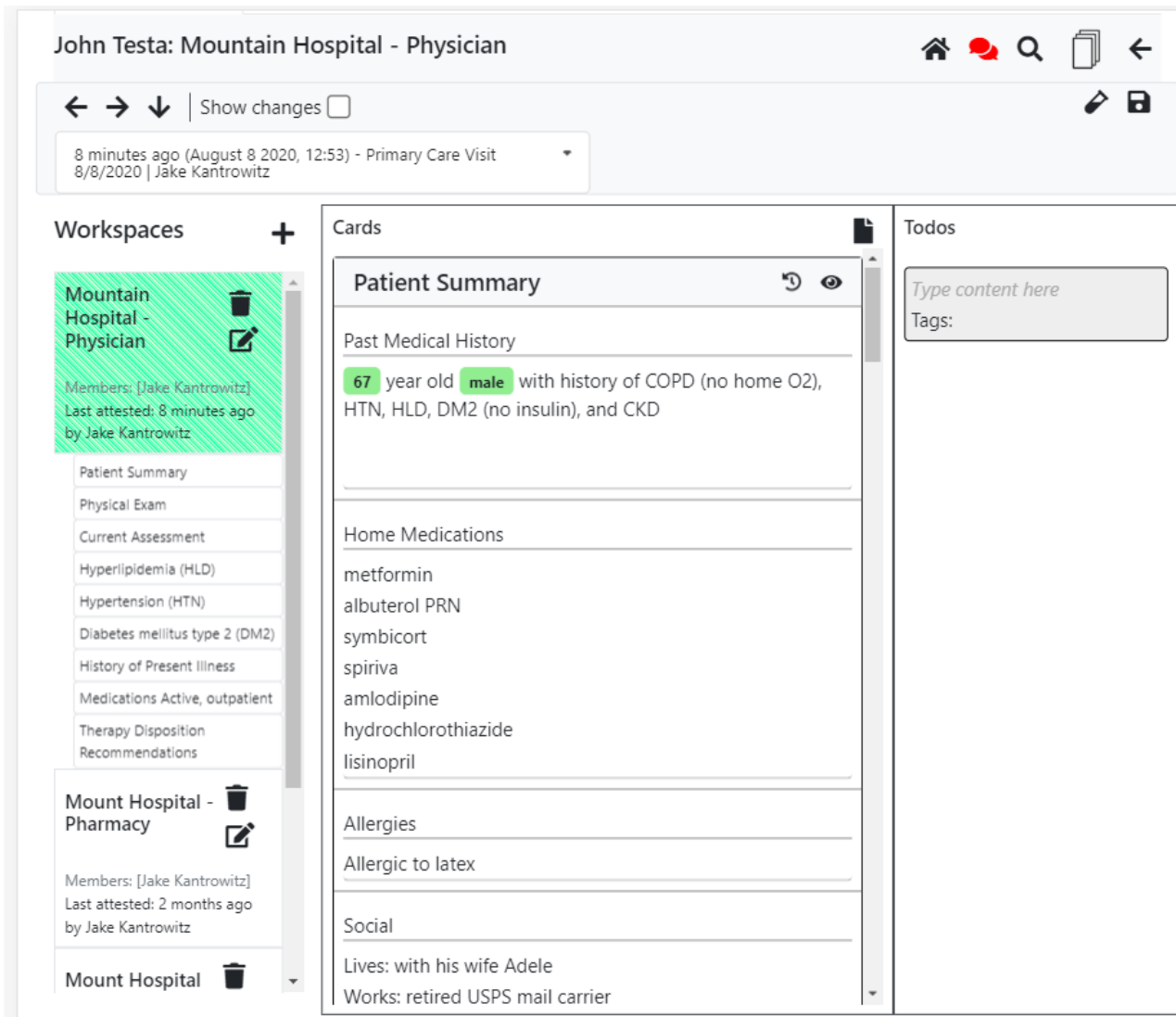


Figure 11. When the show changes button is checked, a track-changes view occurs, enabling later readers to quickly understand what has changed from one state to the next (for instance, the addition of a new diagnosis, medication, and allergy).

The screenshot shows a user interface for a patient's medical record. At the top, it identifies the patient as 'John Testa' at 'Mountain Hospital - Physician'. Below this, there's a navigation bar with a 'Show changes' checkbox that is checked. A dropdown menu shows the last update: '11 minutes ago (August 8 2020, 12:53) - Primary Care Visit' by 'Jake Kantrowitz'. The main interface is divided into three columns: 'Workspaces', 'Cards', and 'Todos'. The 'Workspaces' column lists three workspaces: 'Mountain Hospital - Physician' (with members [Jake Kantrowitz] and last attested 11 minutes ago), 'Mount Hospital - Pharmacy' (with members [Jake Kantrowitz] and last attested 2 months ago), and 'Mount Hospital'. The 'Cards' column shows a 'Patient Summary' card with sections for 'Past Medical History', 'Home Medications', and 'Allergies'. The 'Past Medical History' section highlights '67 year old male with history of COPD (no home O2), HTN, HLD, DM2 (no insulin), and CKD'. The 'Home Medications' section lists 'metformin', 'albuterol PRN', 'symbicort', 'spiriva', 'amlodipine', 'hydrochlorothiazide', and 'lisinopril'. The 'Allergies' section highlights 'NKDA Allergic to latex'. The 'Todos' column is currently empty, with a placeholder 'Type content here' and 'Tags:'.

Team View

Another major source of information scatter rests at the level of the *team* or *patient group*. An extremely common EMR use case involves clinicians managing groups of patients (eg, an inpatient team or an outpatient panel) and either (1) looking for updates in the state of any patient in the group or (2) *running the list*—identifying the status and action items of a group of patients in rapid sequence. Team management and *list-running* are core features of our software. The *team view* interface aggregates the latest state of all workspaces for all patients on a particular team, enabling highly granular actions to be taken for individual patients from a single screen without any requirement to navigate between charts. The team view displays the list of cards for each workspace, enabling quick addition or editing of specific card fields with new single pieces of information. This enables immediate placement of information in topic-appropriate locations—a feature designed to facilitate keeping the chart organized, preventing the need to redocument the same information in different places or scatter information across multiple locations. In addition, the to-do items for each

workspace are aggregated and viewable on one page, enabling efficient list running even for large groups of patients.

Discussion

Principal Findings

We demonstrated the feasibility of building an EMR that does not use the *note* as the core organizational unit for unstructured data. This is made possible by reimagining the chart as a fully collaborative, dynamic workspace, with information organized primarily by topic rather than by time or thread. This paradigm shift facilitates the use of a wide variety of strategies and design elements that have the potential to reduce information overload and information scatter, two of the key pathologies of modern electronic documentation. Our system is designed to accomplish this goal by disincentivizing the behaviors that lead to these pathologies, such as redundant documentation over time and across multiple threads. We believe that this type of EMR data organization has the potential to reduce clinician documentation time; increase direct face-to-face time with patients; mitigate medical errors resulting from conflicting or erroneous information; create cleaner, more intuitive patient charts; and

improve clinician satisfaction with the EMR. We also believe that our organizational paradigm is more well-suited to the actual practice of medicine in the 21st century, with its assumptions of team-based medicine, granular data updates (eg, from SMS text messages, phone calls, biosensors, or digital health devices), and large amounts of information to manage and organize. We hope this study can begin a conversation among clinicians and institutions about the pros and cons of using *notes* as the primary organizational paradigm going forward.

In addition, our system casts doubt on traditional delineations of *structured data* and *unstructured data* by enabling individual clinicians to create their own structured data elements for customized workflows. This decision is motivated by the desire to enable maximal customization for common workflows, such as the workup of common presentations. Too little structure (eg, undifferentiated blocks of free text or loosely templated notes) limits the potential for efficiency gains through automated or stereotyped documentation workflows, whereas too much structure (eg, entering all data in the form of checkboxes, radio buttons, and drop-down menus) is time consuming and implausible for narrative information such as patient histories and clinician thought processes. Rather than being limited by historical decisions about what *structured data* should be (eg, meaningful use requirements or EMR certification guidelines), our system empowers clinicians, teams, or institutions to make their own decisions about how to maximize the gains and minimize the costs of documenting in a structured format.

Our study is a logical extension of the dialog around the problem-oriented medical record, which similarly organizes information primarily by topic. However, we expand the topic-based organization to cover the entirety of medical data in the chart and not just the data that can be represented by *medical problems* or *diagnoses*. In addition, we built the assumptions of full editability and collaboration into the core of the system design. Fully collaborative documentation systems have previously been discussed in different contexts [27], and many studies have pointed out how EMR design paradigms can facilitate or block collaboration between clinicians [28,29]. Our work builds on these discussions, and we believe that fully collaborative documentation systems are a key step in reducing documentation burden and information overload.

We have piloted our system with small user groups on mock patient records but have not yet tested it at scale. In addition, to empirically evaluate whether our system succeeds at its goal of reducing information chaos, direct comparisons between note-based paradigms and non-note-based paradigms will be necessary. This will require further development of EMR usability evaluation frameworks, including a standardized set of metrics for comparing EMRs on relevant end points (including efficiency, cognitive load, and time spent documenting). In future work, we aim to develop such a framework and use it to compare our system with other EMR systems. Another key step in this process is the development of open-source standardized data sets with dummy patient records, designed to facilitate head-to-head comparison of systems at particular common clinical tasks, such as information retrieval, chart summarization, various granularities of data entry, or clinical communication. One could imagine a publicly available EMR "obstacle course" with metrics to quantify the performance of an EMR at common clinical tasks and would enable standardized comparisons for clinicians and institutions looking to reduce EMR time. Such a system could be used not only to evaluate the impact of simple interface changes or feature additions but also to quantify the impact of different organizational paradigms.

Full-fledged EMRs are more than mere documentation interfaces and include functionality to place orders, prescribe medications, and perform population health analyses. These functionalities were not the focus of this study, as we focused primarily on information input and retrieval, as well as task management and collaboration. We believe that non-note interfaces open new possibilities for thinking about order entry, decision support, and other core EMR tasks, as well as how to integrate these tasks with *documentation* per se. In the interim, clinicians could also use such systems as improved *documentation assembly* interfaces in parallel with existing EMRs. In our system, individual workspace states can be exported as raw-text *notes* compatible with EMRs that operate under note-based paradigms. Our system can thus be operationalized either as part of a new EMR or alongside existing EMRs as a separate documentation assembly and information retrieval interface. Ultimately, we hope that clinicians, health systems, and technology vendors will consider the benefits of building and deploying EMRs that operate entirely using collaborative, dynamic, and problem-oriented documentation paradigms.

Acknowledgments

The authors would like to thank Tessa Cook, MD, PhD, and Kalyn Reddy, MD, for helping to think through the early aspects of this work.

Conflicts of Interest

JS, JK, AS, and WB are cofounders of River Records, a health care technology company focused on improving clinician workflows.

Multimedia Appendix 1

Information chaos.

[[DOCX File, 14 KB - formative_v5i11e23789_app1.docx](#)]

Multimedia Appendix 2

Technical details.

[\[DOCX File , 13 KB - formative_v5i11e23789_app2.docx \]](#)**References**

1. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med* 2016 Dec 06;165(11):753-760. [doi: [10.7326/M16-0961](#)] [Medline: [27595430](#)]
2. Beasley JW, Wetterneck TB, Temte J, Lapin JA, Smith P, Rivera-Rodriguez AJ, et al. Information chaos in primary care: implications for physician performance and patient safety. *J Am Board Fam Med* 2011;24(6):745-751 [[FREE Full text](#)] [doi: [10.3122/jabfm.2011.06.100255](#)] [Medline: [22086819](#)]
3. Carayon P, Du S, Brown R, Cartmill R, Johnson M, Wetterneck TB. EHR-related medication errors in two ICUs. *J Healthc Risk Manag* 2017 Jan;36(3):6-15. [doi: [10.1002/jhrm.21259](#)] [Medline: [28099789](#)]
4. Bowman S. Impact of electronic health record systems on information integrity: quality and safety implications. *Perspect Health Inf Manag* 2013 Oct 1;10:1c [[FREE Full text](#)] [Medline: [24159271](#)]
5. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan W, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med* 2017 Sep;15(5):419-426 [[FREE Full text](#)] [doi: [10.1370/afm.2121](#)] [Medline: [28893811](#)]
6. Overhage JM, McCallie D. Physician time spent using the electronic health record during outpatient encounters: a descriptive study. *Ann Intern Med* 2020 Feb 04;172(3):169-174. [doi: [10.7326/M18-3684](#)] [Medline: [31931523](#)]
7. Gardner RL, Cooper E, Haskell J, Harris DA, Poplau S, Kroth PJ, et al. Physician stress and burnout: the impact of health information technology. *J Am Med Inform Assoc* 2019 Feb 01;26(2):106-114 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy145](#)] [Medline: [30517663](#)]
8. Collier R. Electronic health records contributing to physician burnout. *Can Med Asso J* 2017 Nov 13;189(45):1405-1406 [[FREE Full text](#)] [doi: [10.1503/cmaj.109-5522](#)] [Medline: [29133547](#)]
9. Babbott S, Manwell LB, Brown R, Montague E, Williams E, Schwartz M, et al. Electronic medical records and physician stress in primary care: results from the MEMO Study. *J Am Med Inform Assoc* 2014 Mar;21(e1):100-106 [[FREE Full text](#)] [doi: [10.1136/amiainjnl-2013-001875](#)] [Medline: [24005796](#)]
10. Melnick ER, Dyrbye LN, Sinsky CA, Trockel M, West CP, Nedelec L, et al. The association between perceived electronic health record usability and professional burnout among US physicians. *Mayo Clin Proc* 2020 Mar;95(3):476-487 [[FREE Full text](#)] [doi: [10.1016/j.mayocp.2019.09.024](#)] [Medline: [31735343](#)]
11. Chang B, Kanagaraj M, Neely B, Segall N, Huang E. Triangulating methodologies from software, medicine and human factors industries to measure usability and clinical efficacy of medication data visualization in an electronic health record system. *AMIA Jt Summits Transl Sci Proc* 2017 Jul 26;2017:473-482 [[FREE Full text](#)] [Medline: [28815147](#)]
12. Xu L, Wen D, Zhang X, Lei J. Assessing and comparing the usability of chinese EHRs used in two peking university hospitals to EHRs used in the US: a method of RUA. *Int J Med Inform* 2016 May;89:32-42. [doi: [10.1016/j.ijmedinf.2016.02.008](#)] [Medline: [26980357](#)]
13. Saitwal H, Feng X, Walji M, Patel V, Zhang J. Assessing performance of an electronic health record (EHR) using cognitive task analysis. *Int J Med Inform* 2010 Jul;79(7):501-506. [doi: [10.1016/j.ijmedinf.2010.04.001](#)] [Medline: [20452274](#)]
14. Rizvi RF, Marquard JL, Hultman GM, Adam TJ, Harder KA, Melton GB. Usability evaluation of electronic health record system around clinical notes usage-an ethnographic study. *Appl Clin Inform* 2017 Oct;8(4):1095-1105 [[FREE Full text](#)] [doi: [10.4338/ACI-2017-04-RA-0067](#)] [Medline: [29241247](#)]
15. Blecker S, Pandya R, Stork S, Mann D, Kuperman G, Shelley D, et al. Interruptive versus noninterruptive clinical decision support: usability study. *JMIR Hum Factors* 2019 Apr 17;6(2):e12469 [[FREE Full text](#)] [doi: [10.2196/12469](#)] [Medline: [30994460](#)]
16. Belden JL, Koopman RJ, Patil SJ, Lowrance NJ, Petroski GF, Smith JB. Dynamic electronic health record note prototype: seeing more by showing less. *J Am Board Fam Med* 2017;30(6):691-700 [[FREE Full text](#)] [doi: [10.3122/jabfm.2017.06.170028](#)] [Medline: [29180544](#)]
17. Sturmberg JP, Martin CM, Katerndahl DA. Systems and complexity thinking in the general practice literature: an integrative, historical narrative review. *Ann Fam Med* 2014;12(1):66-74 [[FREE Full text](#)] [doi: [10.1370/afm.1593](#)] [Medline: [24445105](#)]
18. Sturmberg JP, Martin CM, Katerndahl DA. It is complicated! - misunderstanding the complexities of 'complex'. *J Eval Clin Pract* 2017 Apr;23(2):426-429. [doi: [10.1111/jep.12579](#)] [Medline: [27307382](#)]
19. Katerndahl DA. Lessons from Jurassic Park: patients as complex adaptive systems. *J Eval Clin Pract* 2009 Aug;15(4):755-760. [doi: [10.1111/j.1365-2753.2009.01228.x](#)] [Medline: [19674231](#)]
20. Smith SW, Koppel R. Healthcare information technology's relativity problems: a typology of how patients' physical reality, clinicians' mental models, and healthcare information technology differ. *J Am Med Inform Assoc* 2014;21(1):117-131 [[FREE Full text](#)] [doi: [10.1136/amiainjnl-2012-001419](#)] [Medline: [23800960](#)]
21. Stephens MB, Gimbel RW, Pangaro L. Commentary: The RIME/EMR Scheme: an educational approach to clinical documentation in electronic medical records. *Acad Med* 2011;86(1):11-14. [doi: [10.1097/acm.0b013e3181ff7271](#)]

22. Weed LL. Medical records that guide and teach. *N Engl J Med* 1968 Mar 14;278(11):593-600. [doi: [10.1056/NEJM196803142781105](https://doi.org/10.1056/NEJM196803142781105)] [Medline: [5637758](https://pubmed.ncbi.nlm.nih.gov/5637758/)]
23. Gorman P, Ash J, Lavelle M, Lyman J, Maier D, Weaver M, et al. Bundles in the wild: managing information to solve problems and maintain situation awareness. *Library Trends*, University of Illinois. 2000. URL: https://www.ideals.illinois.edu/bitstream/handle/2142/8344/librarytrendsv49i2e_opt.pdf?sequence=1&isAllowed=y [accessed 2021-03-07]
24. Graedon T. Is larry weed right? *Society for Participatory Medicine*. 2013. URL: <https://participatorymedicine.org/journal/opinion/editorials/2013/03/18/is-larry-weed-right/> [accessed 2020-12-23]
25. Weed LL, Weed L. Diagnosing diagnostic failure. *Diagnosis (Berl)* 2014 Jan 01;1(1):13-17. [doi: [10.1515/dx-2013-0020](https://doi.org/10.1515/dx-2013-0020)] [Medline: [29539981](https://pubmed.ncbi.nlm.nih.gov/29539981/)]
26. Kleppmann M, Wiggins A, van HP, McGranaghan M. Local-first software: you own your data, in spite of the cloud. In: *Proceedings of the ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software*. 2019 Presented at: SPLASH '19: 2019 ACM SIGPLAN International Conference on Systems, Programming, Languages, and Applications: Software for Humanity; October 23 - 24, 2019; New York, NY p. 154-178. [doi: [10.1145/3359591.3359737](https://doi.org/10.1145/3359591.3359737)]
27. Warner JL, Smith J, Wright A. It's time to wikify clinical documentation: how collaborative authorship can reduce the burden and improve the quality of the electronic health record. *Acad Med* 2019 May;94(5):645-650. [doi: [10.1097/ACM.0000000000002613](https://doi.org/10.1097/ACM.0000000000002613)] [Medline: [30681451](https://pubmed.ncbi.nlm.nih.gov/30681451/)]
28. Vos JF, Boonstra A, Kooistra A, Seelen M, van Offenbeek M. The influence of electronic health record use on collaboration among medical specialties. *BMC Health Serv Res* 2020 Jul 22;20(1):676 [FREE Full text] [doi: [10.1186/s12913-020-05542-6](https://doi.org/10.1186/s12913-020-05542-6)] [Medline: [32698807](https://pubmed.ncbi.nlm.nih.gov/32698807/)]
29. Rashotte J, Varpio L, Day K, Kuziemy C, Parush A, Elliott-Miller P, et al. Mapping communication spaces: the development and use of a tool for analyzing the impact of EHRs on interprofessional collaborative practice. *Int J Med Inform* 2016 Sep;93:2-13. [doi: [10.1016/j.ijmedinf.2016.05.003](https://doi.org/10.1016/j.ijmedinf.2016.05.003)] [Medline: [27435942](https://pubmed.ncbi.nlm.nih.gov/27435942/)]

Abbreviations

CDS: clinical decision support

CHF: congestive heart failure

EMR: electronic medical record

RIME: reporter, interpreter, manager, and educator

Edited by G Eysenbach, R Kukafka; submitted 23.08.20; peer-reviewed by SK Mukhiya, D Chartash; comments to author 30.10.20; revised version received 24.12.20; accepted 20.09.21; published 09.11.21.

Please cite as:

Steinkamp J, Sharma A, Bala W, Kantrowitz JJ

A Fully Collaborative, Noteless Electronic Medical Record Designed to Minimize Information Chaos: Software Design and Feasibility Study

JMIR Form Res 2021;5(11):e23789

URL: <https://formative.jmir.org/2021/11/e23789>

doi: [10.2196/23789](https://doi.org/10.2196/23789)

PMID: [34751651](https://pubmed.ncbi.nlm.nih.gov/34751651/)

©Jackson Steinkamp, Abhinav Sharma, Wasif Bala, Jacob J Kantrowitz. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 09.11.2021. 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

Home-Based Exercise Program for Patients With Combined Advanced Chronic Cardiac and Pulmonary Diseases: Exploratory Study

Cyrille Herkert¹, MD; Lidwien Graat-Verboom², MD, PhD; Judith Gilsing-Fernhout³, BHLth; Manon Schols³, BHLth; Hareld Marijn Clemens Kemps^{1,4}, MD, PhD

¹Department of Cardiology, Máxima Medical Center, Eindhoven, Netherlands

²Department of Pulmonology, Máxima Medical Center, Eindhoven, Netherlands

³ParaMáx: Center for Paramedic Care, Máxima Medical Center, Eindhoven, Netherlands

⁴Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands

Corresponding Author:

Cyrille Herkert, MD

Department of Cardiology

Máxima Medical Center

Dominee Theodor Flidnerstraat 1

Eindhoven, 5631 BM

Netherlands

Phone: 31 408888220

Fax: 31 408885275

Email: cyrille.herkert@mmc.nl

Abstract

Background: As chronic cardiac and pulmonary diseases often coexist, there is a need for combined physical home-based rehabilitation programs, specifically addressing older patients with advanced disease stages.

Objective: The primary aim of this study is to evaluate the completion and adherence rates of an 8-week, home-based exercise program for patients with advanced cardiopulmonary disease. The secondary end points include patient satisfaction; adverse events; and program efficacy in terms of change in functional capacity, level of dyspnea, and health-related quality of life.

Methods: The participants received a goal-oriented, home-based exercise program, and they used a wrist-worn activity tracker to record their exercise sessions. Activity tracker data were made visible on a digital platform, which was also equipped with several other features such as short instruction videos on how to perform specific exercises. The participants received weekly coaching by a physiotherapist and an occupational therapist through video communication.

Results: In all, 10 patients with advanced combined cardiopulmonary disease participated (median age 71, IQR 63-75 years), and 50% (5/10) were men. Of the 10 participants, 9 (90%) completed the 8-week program. Median adherence to the exercise prescription was 75% (IQR 37%-88%), but it declined significantly when the program was divided into 2-week periods (first 2 weeks: 86%, IQR 51%-100%, and final 2 weeks: 57%, IQR 8%-75%; $P=.03$). The participants were highly satisfied with the program (Client Satisfaction Questionnaire: median score 29, IQR 26-32, and Purpose-Designed Questionnaire: median score 103, IQR 92-108); however, of the 9 participants, 4 (44%) experienced technical issues. The Patient-Specific Complaints Instrument scores declined, indicating functional improvement (from median 7.5, IQR 6.1-8.9, to median 5.7, IQR 3.8-6.7; $P=.01$). Other program efficacy metrics showed a trend toward improvement.

Conclusions: Home-based cardiopulmonary telerehabilitation for patients with severe combined cardiopulmonary disease is feasible in terms of high completion and satisfaction rates. Nevertheless, a decrease in adherence during the program was observed, and some of the participants reported difficulties with the technology, indicating the importance of the integration of behavior change techniques, using appropriate technology.

Trial Registration: Netherlands Trial Register NL9182; <https://www.trialregister.nl/trial/9182>

(*JMIR Form Res* 2021;5(11):e28634) doi:[10.2196/28634](https://doi.org/10.2196/28634)

KEYWORDS

home-based exercise; cardiac diseases; pulmonary diseases; comorbidities; elderly

Introduction

Background

Chronic cardiac and pulmonary diseases often coexist because they share similar risk factors such as older age, cigarette smoking, sedentary lifestyle, and persistent low-grade systemic inflammation [1]. The prevalence of cardiac diseases such as ischemic heart disease, chronic heart failure (CHF), and atrial fibrillation is higher in patients with chronic obstructive pulmonary disease (COPD) than in the general population (30%, 20%, and 13%, respectively) [1,2]. In addition, in patients with CHF and ischemic heart disease, COPD and asthma are more prevalent compared with patients without cardiovascular diseases [3]. As chronic cardiac and pulmonary diseases frequently overlap in symptoms and clinical course and share a high morbidity, with recurrent hospital admissions, there is a need for an integrated management of these often older and frail patients. Ideally, this applies not only to diagnostic and therapeutic regimens but also to physical rehabilitation.

Exercise-based cardiac rehabilitation (CR) and pulmonary rehabilitation (PR) are highly recommended for a wide range of cardiac and respiratory disorders because they result in improvement of exercise capacity and quality of life (QoL) and reduction in hospital admissions related to the disease [4,5]. Despite these positive effects as well as international guidelines recommending CR and PR, participation remains low. Concerning CR, approximately 30% of all eligible patients attend a CR program [6], with the lowest uptake in patients with chronic cardiovascular diseases such as CHF [7]. This underuse is attributed not only to low referral rates but also to patient-related factors such as older age, comorbidities, and larger distance to the nearest rehabilitation center [8]. As in the case of CR, PR uptake in patients with COPD is low [9]. Lack of transport, timing of the program, and burden of illness are major factors impeding PR attendance [10]. To overcome these barriers, there is an increasing interest in home-based rehabilitation programs. Home-based CR has been shown to have effects comparable with those of center-based CR on exercise capacity, health-related QoL (HRQoL), and mortality in the short term, both in patients diagnosed with coronary artery disease and in patients with CHF, with a higher compliance in home-based CR [11]. The results of a Cochrane review on telerehabilitation in patients with chronic respiratory diseases showed that there is a limited number of studies available that provide evidence for a similar outcome compared with traditional PR [12].

To date, there are 2 major gaps in evidence with respect to cardiac and pulmonary telerehabilitation. First, the studies on cardiac and pulmonary telerehabilitation mainly focus on patients classified as low risk, that is, younger patients, without major comorbidities, in a stable phase of the disease, rather than on the merging group of older and frail patients. These vulnerable patients, however, might benefit most from rehabilitation because of impairment in multiple domains of

physical functioning, which lead to recurrent hospital admissions [13]. In fact, a pilot study on rehabilitation in older patients with acute decompensated heart failure showed promising results in improving physical function and reducing rehospitalization [13]. Although study retention and adherence were acceptable in this center-based rehabilitation program, offering home-based programs for these patients might be even more successful to overcome the aforementioned barriers to center-based rehabilitation, which are particularly common in this patient category. Second, cardiac and pulmonary disease management are seldom integrated. To our knowledge, there is only one trial focusing on telerehabilitation of patients with combined CHF and COPD, showing that their combined telerehabilitation program is feasible, safe, and effective [14]. As this program focused on patients with middle-severe combined CHF and COPD, the effectivity in patients with advanced combined cardiopulmonary disease remains to be established.

Objective

To address the current gaps in telerehabilitation, this exploratory study was designed to assess the feasibility of a personalized, home-based, goal-oriented exercise program in patients with advanced, combined chronic cardiac and pulmonary diseases. The results of this study will be used to design a future, large clinical trial.

Methods

Study Design and Population

This is a single-center feasibility study designed to explore program completion and adherence, as well as possible barriers to adherence, of a novel, personalized, home-based telerehabilitation program for patients with combined advanced chronic cardiac and pulmonary diseases. The secondary end points include patient satisfaction; adverse events; and an exploratory analysis of program efficacy in terms of functional capacity, level of dyspnea, and HRQoL. The results of this study will be used to guide the design of a larger clinical trial. The study was performed at the Máxima Medical Center in Veldhoven, Netherlands. In all, 10 patients diagnosed with a combination of chronic cardiac and pulmonary diseases were included. The eligible patients were randomly selected from a group of patients who were participating in a telemonitoring program for combined chronic cardiac and pulmonary disease because of 1 or more hospital admissions within the past year on account of a cardiac or pulmonary exacerbation or frequent exacerbations treated at the outpatient clinic. The exclusion criteria were neurological, orthopedic, or peripheral vascular conditions preventing the patient from performing exercise; hemodynamic significant valvular disease; and proven cardiac ischemia or heart rhythm disturbances at a low-intensity exercise level. All patients provided written informed consent. The study was approved by the local medical ethical committee of Máxima Medical Center and was conducted in accordance with the declaration of Helsinki. The study was registered in the Netherlands Trial Register (NL9182).

Outcome Measures

Program Completion and Adherence

The primary end points—program completion and adherence—were defined as the percentage of participants who completed the 8-week exercise program and the percentage of completed exercise sessions after receiving a prescription, respectively. To assess adherence, the participants were asked to record their exercise sessions with a wrist-worn activity tracker (Galaxy Watch Active; Samsung Electronics; see *Digital Platform and Activity Tracker* section). On the basis of the study population, characterized by a high disease severity and the presence of several comorbidities, we acknowledge that program completion and adherence are challenging. Nevertheless, given the fact that the entire program was highly personalized, performed in the patients' home environment, and supported by fairly easy-to-use technology, we hypothesized that most of the participants would be able to complete the program and that adherence to the exercise prescription would be at least 70%.

Patient Satisfaction

Patient satisfaction and possible barriers to adherence were assessed with the validated Client Satisfaction Questionnaire-8 (CSQ-8) [15]. This self-administered questionnaire consists of 8 items, scored on 4-point Likert-type scales. For every patient, a total satisfaction score is calculated with a minimum score of 8 and a maximum score of 32, where higher scores indicate a higher level of satisfaction [16].

Moreover, a self-administered, Purpose-Designed Questionnaire (PDQ) was completed by each participant ([Multimedia Appendix 1](#)). This questionnaire consists of 26 items categorized into 5 topics: provision of information, contact with the therapists, safety, digital platform and activity tracker, and treatment results. Of the 26 items, 23 are scored on a 5-point Likert scale, resulting in a minimum score of 23 and a maximum score of 115 (a higher score indicates a higher level of satisfaction); 1 item is answered with yes or no; 1 item requires the participant to grade (from 0=very bad to 10=excellent) the program in general; and 1 item is an open question seeking comments or suggestions about the program.

Program Efficacy

Subjective program efficacy was evaluated through the following 3 questionnaires or rating scales, which were administered at baseline and again at the end of the program:

- Patient-Specific Complaints (PSC) Instrument. This tool asks the patient to mention 2-3 activities that are difficult to perform because of cardiopulmonary complaints. Each activity is then rated on a scale from 0 (very easy to perform the activity) to 10 (impossible to perform the activity). A mean PSC score is calculated by dividing the total PSC score by the number of chosen activities.
- Modified Medical Research Council (mMRC) Dyspnea Scale [17], which measures the degree of dyspnea severity during daily life activities. This self-administered tool rates the degree of disability in daily life caused by breathlessness on a scale from 0 to 4, where 0 represents not troubled by breathlessness except during strenuous exercise, and 4

represents too breathless to leave the house or breathless when dressing or undressing.

- Self-administered EuroQoL 5-Dimension, 5-Level (EQ-5D-5L) Questionnaire, which measures HRQoL and which consists of 5 questions in 5 domains (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) [18]. Each question is answered on a 5-point severity scale. Moreover, the questionnaire comprises a visual analog scale (from 0=worst imagined health to 100=best imagined health) to grade HRQoL.

The objective measures to evaluate program efficacy, extracted at baseline and again at the end of the program were as follows:

- Number of repetitions during the 1-minute sit-to-stand (1-MSTS) test, which is a valid and responsive tool for measuring exercise capacity [19]. The patients were asked to sit on a chair, which was stabilized against the wall, with their arms crossed over their chest, their knees in an approximate 90-degree angle, and feet flat on the floor. They were asked to perform as many sit-to-stand cycles as possible in 1 minute. The patients were asked to fully stand up and then sit back again, touching the surface of the seat but not the backrest. They were not encouraged during the test but were informed when there were 10 seconds left, timed with a stopwatch.
- Handgrip strength as a measure of muscle strength, measured with the Jamar hydraulic hand dynamometer (JLW Instruments) [20]. The patients were asked to be seated, with shoulders in neutral position, elbows flexed at 90 degrees, and wrists in neutral position with the thumb facing upward. The forearm was not to rest on the arm of the chair. Three trials of handgrip strength were performed using each hand, beginning with the dominant hand. The patients were encouraged to squeeze as hard as they could. The maximum grip strength from all 6 trials was used.

Exercise Program

The exercise program started with a combined intake supervised by both the physiotherapist and occupational therapist at the outpatient physiotherapy clinic of Máxima Medical Center. The intake consisted of an assessment of the daily life activities of each participant and the perceived barriers. On the basis of this interview, 2 to 3 problematic activities were defined that the participant wished to improve. These activities were rated with the PSC Instrument and were translated into patient-specific rehabilitation goals. In addition, the participants were instructed to complete the EQ-5D-5L Questionnaire and the mMRC Dyspnea Scale and to perform the 1-MSTS and handgrip tests.

The 8-week home-based exercise program consisted of a combination of endurance and strength exercise training tailored to the participants' preferences in exercise modality (ie, walking, cycling, or swimming) and availability of training equipment at home. Exercises for (respiratory) muscle strengthening, breathing techniques, and techniques for mobilization of sputum were provided through instruction videos on the digital platform. The amount, duration, and content of the exercise sessions was determined individually by the physiotherapist. The intensity of the prescribed exercise sessions was guided by the rate of perceived exertion (Borg Rating of Perceived Exertion 6-20

Scale), which was filled in on the digital platform by the participant after a completed exercise session. The exercise sessions were recorded by the participant with a wrist-worn activity tracker (see *Digital Platform and Activity Tracker* section).

The occupational therapist primarily focused on the execution of daily life activities and energy balance. For this purpose, the digital platform was supplied with an activity diary to be filled in by the participant, which could be rated afterward by the occupational therapist.

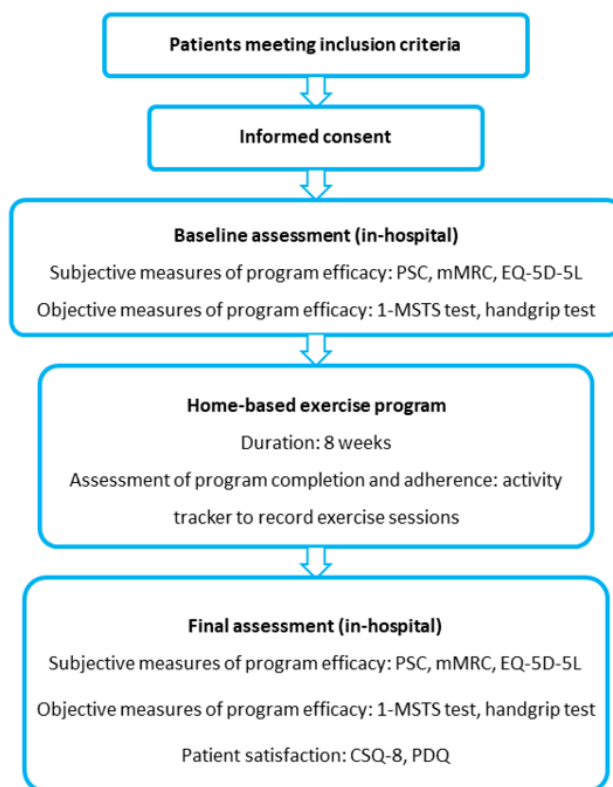
The participants had a weekly video consultation through the digital platform with the physiotherapist or occupational therapist to discuss their progress and to adjust the training scheme if needed. However, the therapists had the option to

adjust the frequency of the video consultations based on shared decision-making with the participant.

At the end of the program, an evaluation took place at the outpatient clinic, again with both therapists. The participants' personal goals were evaluated through the PSC Instrument. In addition, the 1-MSTS and handgrip tests were administered again, along with the EQ-5D-5L Questionnaire and mMRC Dyspnea Scale. Finally, each participant answered 2 questionnaires (CSQ-8 and PDQ) on their satisfaction with the program, and the responses were analyzed after all participants completed the study.

An overview of the participant timeline and study measurements is presented in [Figure 1](#).

Figure 1. Participant timeline and study measurements. 1-MSTS: 1-minute sit-to-stand; CSQ-8: Client Satisfaction Questionnaire-8; EQ-5D-5L: EuroQol 5-Dimension, 5-Level Questionnaire; mMRC: modified Medical Research Council; PDQ: Purpose-Designed Questionnaire; PSC: Patient-Specific Complaints.



Digital Platform and Activity Tracker

This study used a secure digital platform (Mibida BV). The eligible participants were already using this platform for telemonitoring of their cardiopulmonary disease. For the participants in this study, the platform was equipped with the following additional functionalities:

- An overview of the executed exercise sessions (recorded with the activity tracker). For each session, the following metrics were visualized: type of exercise, date of the exercise session, starting time of the session, duration of the session, heart rate course, and perceived Borg score. Actual feedback on these sessions was given by the therapists during the weekly videocall sessions.

- Short instruction films about exercises for (respiratory) muscle strengthening, breathing techniques, and techniques for mobilization of sputum. These films were recorded in advance in collaboration with the physiotherapy department. For each patient, a personalized set of films was selected by the physiotherapist and transferred to the digital platform.
- An activity diary for 1 or more days to be filled in by the participant, which could be scored afterward by the occupational therapist. These scores, represented as numbers, were discussed during the participants' video consultations with the occupational therapist.
- A graphical overview of treatment goals based on the PSC list, with the associated PSC scores. The PSC scores could be filled in periodically by the participant (as was asked by the therapists) to evaluate progress.

- A videocall module for weekly communication with the therapists.
- A tab with reports created by the therapists. Tasks and assignments for the upcoming week were summarized here, as well as progress up to that point.

Exercise sessions were recorded with a wrist-worn activity tracker (Galaxy Watch Active). The watch was mainly used for evaluation of adherence, rather than for adjusting exercise prescription based on specific recorded metrics. The exercise recordings were automatically uploaded to the digital platform through the home Wi-Fi network of the patient.

Statistical Analysis

For this study, we chose a convenience sample of 10 participants, rather than a calculation-based sample size. Nevertheless, we included a very homogeneous group of participants (older adults with a combined cardiopulmonary diagnosis and a recent exacerbation). For this reason, we considered a sample size of 10 patients sufficient to (1) explore whether the approach is feasible to execute on a larger scale and (2) obtain sufficient representative results to enable us to improve the intervention before upscaling. Descriptive statistics were used to describe the studied population with regard to the baseline characteristics and data on adherence, patient satisfaction, and program efficacy. Because of the small patient population, nonparametric tests were used. The Wilcoxon signed-rank test was used to compare differences in the PSC,

mMRC, HRQoL, and 1-MSTS scores, as well as in handgrip strength, before and after the home-based exercise program, and the Friedman test was used to determine whether the change in adherence during the program was significant. Statistical analysis was performed using SPSS software (version 22; IBM Corp). For the analysis, the significance level was set at $P < .05$.

Results

Baseline Characteristics

The aim was to include 10 patients. Therefore, 13 patients were asked to participate, of whom 3 (23%) declined. Of these 3 patients, 2 (67%) provided reasons for declining to participate: (1) starting an extensive treatment for a comorbid disorder and not wanting to combine it with study participation and (2) feeling insecure about performing exercise that would not be directly supervised. The study participants were included between October 2019 and October 2020, had a median age of 71 (IQR 63-75) years, and 50% (5/10) were men. Additional patient baseline characteristics are presented in Table 1, and baseline cardiopulmonary patient profiles are presented in Table 2. Of the 10 participants, 9 (90%) were already familiar with using the digital platform for telemonitoring purposes—the median duration of the use of this platform was 20 (IQR 4-24) months—whereas 1 (10%) participant had never used the digital platform previously because telemonitoring had not started yet for this participant.

Table 1. Baseline characteristics (N=10).

Characteristics	Values
Age (years), median (IQR)	71 (63-75)
Gender, n (%)	
Male	5 (50)
Female	5 (50)
Height (m), median (IQR)	1.68 (1.60-1.84)
Weight (kg), median (IQR)	86 (71-96)
BMI (kg/m ²), median (IQR)	28.2 (27.4-32.1)
Number of comorbidities ^a , median (IQR)	3.0 (0.8-4.3)
Cardiac and pulmonary exacerbations ^b , year before inclusion, median (IQR)	3.5 (1.8-6.0)
LTOT ^c , n (%)	6 (60)
mMRC^d Dyspnea Scale, n (%)	
0	0 (0)
1	1 (10)
2	1 (10)
3	4 (40)
4	4 (40)

^aChronic diseases other than cardiopulmonary diagnosis.

^bEvents characterized by worsening of respiratory symptoms and/or peripheral edema beyond normal day-to-day variation that led to change in medication (ie, diuretics, inhalers, systemic corticosteroids, antibiotics, or a combination). Events might be accompanied by a hospital admission.

^cLTOT: long-term oxygen therapy.

^dmMRC: modified Medical Research Council.

Table 2. Baseline cardiopulmonary patient profiles (N=10).

Patient number	Cardiac diagnosis	LVEF ^a (%)	Pulmonary diagnosis	DLCO ^b
1	HFpEF ^c	63	COPD ^d GOLD ^e IV, group A	43
2	HFrEF ^f	35	COPD GOLD IV, group D	40
3	HFmrEF ^g	47	COPD GOLD II, group B	53
4	HFpEF	53	COPD GOLD IV, group D; recent COVID-19 infection	25
5	HFpEF	60	Asthma, bronchiectasis	— ^h
6	HFpEF	63	COPD GOLD II, group D	29
7	HFrEF	35	Asthma, unilateral emphysema	50
8	HFmrEF	40	COPD GOLD IV, group D	31
9	HFrEF	35	COPD GOLD II, group B; recent COVID-19 infection	68
10	HFpEF	65	COPD GOLD IV, group D	26

^aLVEF: left ventricular ejection fraction.

^bDLCO: diffusing capacity for carbon monoxide.

^cHFpEF: heart failure with preserved ejection fraction.

^dCOPD: chronic obstructive pulmonary disease.

^eGOLD: Global Initiative for Chronic Obstructive Lung Disease.

^fHFrEF: heart failure with reduced ejection fraction.

^gHFmrEF: heart failure with midrange ejection fraction.

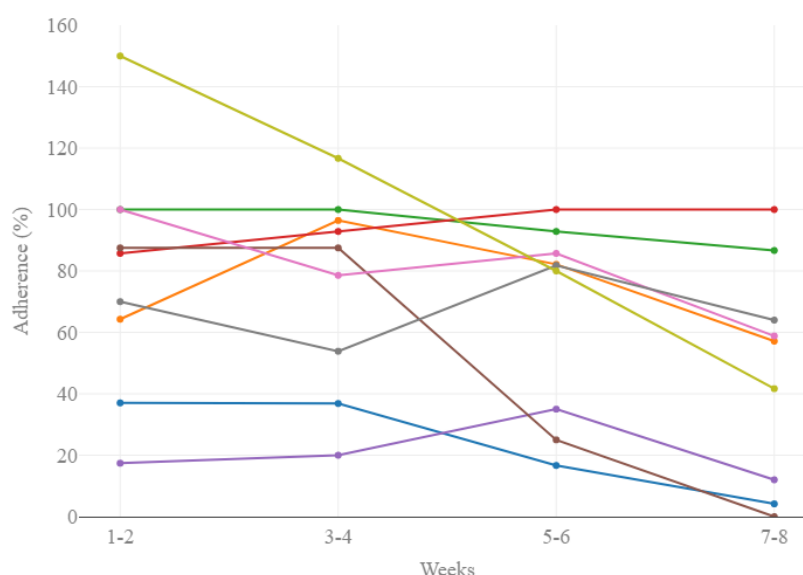
^hNot available.

Study Completion and Adherence

Of the 10 participants, 9 (90%) completed the 8-week exercise program, whereas 1 (10%) participant ended the program prematurely because of a hospital admission (neither related to cardiac or pulmonary disease, nor to study participation) shortly

after inclusion and inability to restart the program after discharge. The median adherence (ie, the percentage of completed exercise sessions after receiving the prescription) over 8 weeks was 75% (IQR 37%-88%), ranging from 20% to 95% (Figure 2).

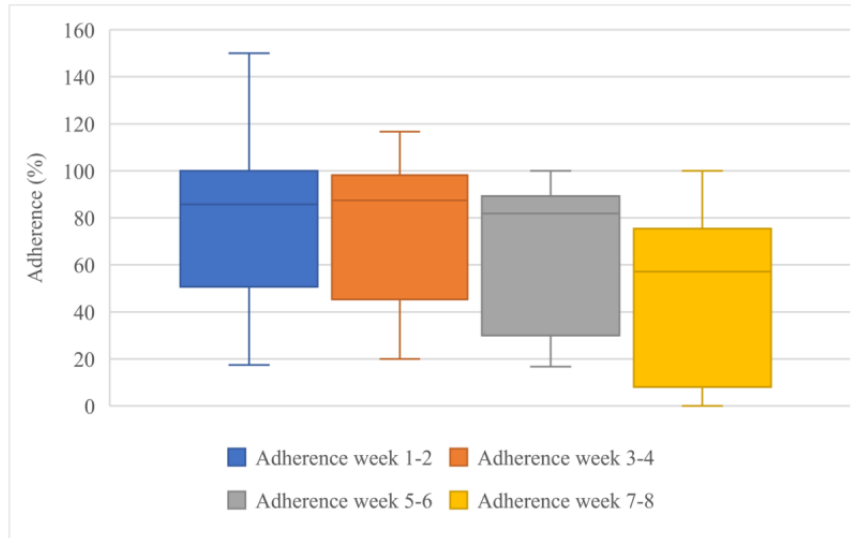
Figure 2. Interpersonal variability in adherence (n=9). Each colored line represents 1 participant.



When we divided the program into periods of 2 weeks, the median adherence decreased significantly: from 86% (IQR 51%-100%) in the first 2 weeks of the program to 57% (IQR

8%-75%) in the final 2 weeks of the program ($P=.03$; Figure 3).

Figure 3. Boxplots showing median adherence over periods of 2 weeks (n=9).



Adverse Events

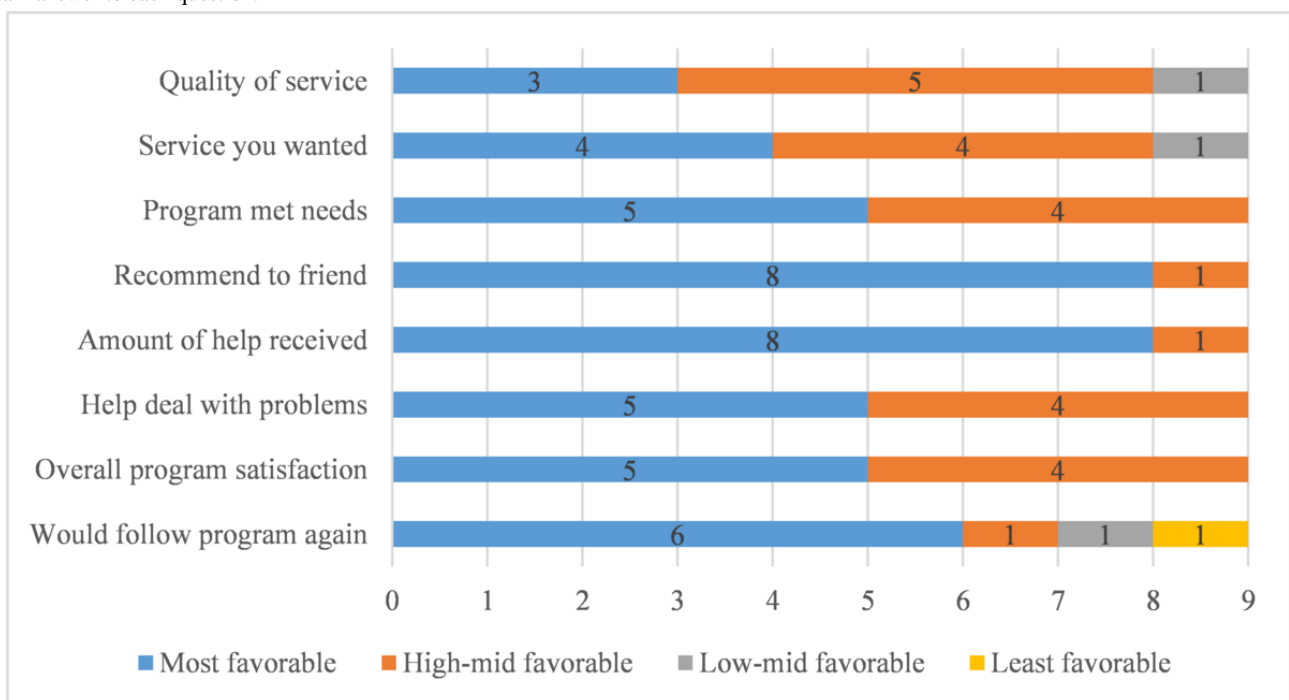
During the study period, no adverse events occurred that were related to the home-based exercise program. Of the 9 participants who completed the program, in the case of 1 (11%) participant, the study was interrupted for several weeks by a severe pulmonary exacerbation in combination with a COVID-19 infection. This participant successfully restarted the study after recovery.

Patient Satisfaction

CSQ-8 Score

The median total CSQ-8 score was 29 (IQR 26-32), which indicates a high level of satisfaction. The total CSQ-8 score ranged between a minimum of 21 and a maximum of 32. Figure 4 illustrates the distribution of each item of the CSQ-8, showing that CSQ questions 4 and 5 were scored best: 89% (8/9) of the participants were highly satisfied with the amount of help received and would definitely recommend the program to a friend, 67% (6/9) would definitely follow the program again, whereas of the 9 participants who completed the program, 1 (11%) participant would definitely not follow it again.

Figure 4. Distribution of each item of the Client Satisfaction Questionnaire-8 (n=9). Data are presented as the number of participants who provided a certain answer to each question.



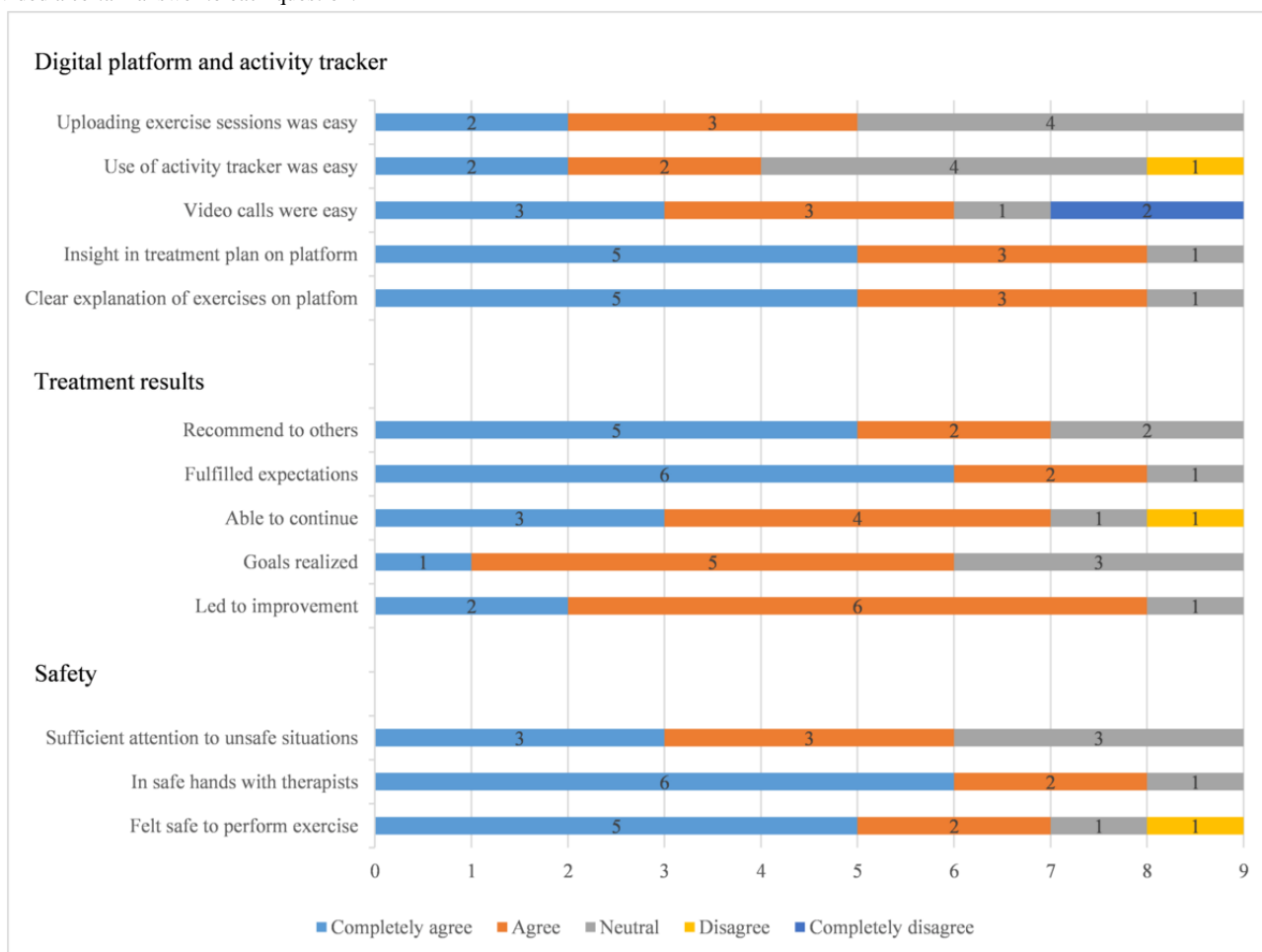
PDQ Score

The median total PDQ score was 103 (IQR 92-108), indicating a high level of satisfaction. The total PDQ score ranged between a minimum of 71 and a maximum of 113.

The subset of items in the topics provision of information (3 questions) and contact with the therapists (7 questions) were responded to with agree or completely agree by 6 to 9 of the participants, indicating high levels of satisfaction. No participant disagreed with these items. Figure 5 shows that the remainder of the topics (digital platform and activity tracker, treatment results, and safety) had a slightly broader distribution. Generally,

agreement with the items of these topics was acceptable; however, regarding the digital platform and activity tracker, of the 9 participants who completed the program, 2 (22%) completely disagreed with the statement that making video calls was easy and 1 (11%) disagreed with the statement that the use of the activity tracker was easy, whereas for the topic treatment results, 1 (11%) disagreed with the statement about being able to continue the exercises after the conclusion of the program. Regarding safety, of the 9 participants, 7 (78%) completely agreed or agreed with the statement regarding whether they felt it was safe to perform the exercises at home, 1 (11%) chose the neutral response, and 1 (11%) disagreed.

Figure 5. Distribution of each item in 3 topics of the Purpose-Designed Questionnaire (n=9). Data are presented as the number of participants who provided a certain answer to each question.



The participants gave the program a median grade of 9.0 (IQR 8.0-9.5), where a grade of 0 represents very bad and 10 represents excellent. The final question of the PDQ is an open one, which gives participants the option to write down comments. Of the 9 participants, 6 (67%) answered this question. The most frequently reported comments were experiencing trouble making a proper video connection with the therapists (3/9, 33%) and experiencing trouble with recording the exercise sessions with the activity tracker (4/9, 44%). Of the 9 participants, 4 (44%) answered yes to the question regarding whether they experienced technical issues to a certain extent.

Program Efficacy

Subjective Measures

The PSC score significantly decreased after completion of the program, from median 7.5 (IQR 6.1-8.9) to median 5.7 (IQR 3.8-6.7; $P=.01$). The mMRC dyspnea score decreased nonsignificantly after completion of the program, from median 3.0 (IQR 2.8-4.0) to median 2.0 (IQR 1-3; $P=.07$). The HRQoL score showed some trends toward improvement; however, no statistically significant changes were observed (Table 3).

Table 3. Health-related quality of life (measured with the self-administered EuroQol 5-Dimension, 5-Level Questionnaire) before and after the program (N=9).

	Baseline	After program completion	<i>P</i> value
Mobility, n (%)			.32
No problems	1 (11)	0 (0)	
Slight problems	0 (0)	2 (22)	
Moderate problems	5 (56)	6 (67)	
Severe problems	3 (33)	1 (11)	
Unable to walk	0 (0)	0 (0)	
Self-care, n (%)			.99
No problems	4 (44)	4 (44)	
Slight problems	3 (33)	2 (22)	
Moderate problems	0 (0)	2 (22)	
Severe problems	2 (22)	1 (11)	
Unable to wash or dress	0 (0)	0 (0)	
Usual activities, n (%)			.16
No problems	2 (22)	2 (22)	
Slight problems	1 (11)	3 (33)	
Moderate problems	2 (22)	3 (33)	
Severe problems	4 (44)	1 (11.1)	
Unable to do usual activities	0 (0)	0 (0)	
Pain or discomfort, n (%)			.05
No pain or discomfort	1 (11)	4 (44)	
Slight pain or discomfort	6 (67)	4 (44)	
Moderate pain or discomfort	2 (22)	1 (11)	
Severe pain or discomfort	0 (0)	0 (0)	
Extreme pain or discomfort	0 (0)	0 (0)	
Anxiety or depression, n (%)			.56
Not anxious or depressed	5 (56)	7 (78)	
Slightly anxious or depressed	3 (33)	0 (0)	
Moderately anxious or depressed	1 (11)	2 (22)	
Severely anxious or depressed	0 (0)	0 (0)	
Extremely anxious or depressed	0 (0)	0 (0)	
VAS ^a , median (IQR)	50.0 (42.5-70.0)	60.0 (50.0-75.0)	.09

^aVAS: visual analog scale.

Objective Measures

The number of repetitions during the 1-MSTS test showed a nonsignificant increase after completion of the program, from a median of 15 (IQR 12-19) repetitions to a median of 16 (IQR 13-21) repetitions ($P=.05$). In addition, no significant differences in handgrip strength were observed, from a median of 30 (IQR 25-34) kg to a median of 30 (IQR 23-33) kg ($P=.73$).

Financial Costs

The financial costs of the program can be divided into (1) costs with regard to the deployment of health care professionals and (2) costs with regard to technology (activity tracker, digital platform, and laptops that were needed for the therapists to perform video consultations). An overview of the intervention costs per participant is presented in [Table 4](#).

Table 4. Intervention costs per participant.

Type of costs	Costs per participant €(US \$)
Health care professionals	
Intake with 2 therapists (1 hour)	79.94 (92.69)
Video consultations (6×15 minutes)	59.96 (69.53)
Final evaluation with 2 therapists (30 minutes)	39.97 (46.35)
Activity trackers (2 months)	8.33 (9.66)
Hosting costs of digital platform (2 months) ^a	13 (15)
Laptops (2 months) ^a	0.31 (0.36)
Total costs per participant per program	201.51 (233.66)

^aBased on an annual number of 200 participants in regular care setting.

Discussion

Overview

This exploratory study showed that home-based cardiopulmonary telerehabilitation for patients with severe combined cardiac and pulmonary disease is feasible in terms of program completion and is associated with high satisfaction rates and functional improvement. However, a decrease in adherence during the program was observed, and some participants reported difficulties with the technology.

Program Completion and Adherence

Our study showed that of the 10 participants, 9 (90%) completed the program, but the median adherence was lower (75%; (IQR 37%-88%). It should be noted, however, that the actual adherence may have been underestimated because we based our findings solely on the number of recorded sessions with the activity tracker, which was experienced as troublesome by some patients. Other studies evaluating home-based CR in patients with CHF generally showed a high mean adherence but with a wide variation ranging from 54% to 110% [21]. In patients with COPD too, adherence to home-based exercise programs was highly variable, ranging from 21% [22] to 93.5% [23]. The study by Bernocchi et al [14] reported both a high completion rate of 86% and high adherence of 93% during a 4-month home-based rehabilitation program for patients with middle-severe combined CHF and COPD. However, because adherence was monitored by means of a written physical activity diary compared with sensor-based registration in our study, it is difficult to directly compare these results. Other factors explaining the broad variation in the adherence rates in home-based programs include the design of the program (eg, directly supervised home-based sessions vs unsupervised sessions); the number of contact moments with the telerehabilitation institution; the use of motivational strategies during the program; and the population being studied with regard to age and frailty, disease severity, clinical stability of the disease, and presence of comorbidities.

Although completion of the telerehabilitation program was high, we observed a decrease in adherence to the prescribed exercise sessions during the study period. A possible explanation for reduced adherence is temporary worsening of cardiopulmonary

physical complaints or other coexisting diseases. Although the patients were also participating in a telemonitoring program focused on symptoms and vital signs, an actual integration of telemonitoring and the home-based exercise program was lacking. Ideally, when worsening of physical complaints was detected by telemonitoring, the rehabilitation therapists should have been informed immediately to adjust the exercise program to the patients' physical capacity. Indeed, the study by Loeckx et al [24] showed that a lack in program adjustment during an exacerbation of COPD was experienced as a barrier by patients with COPD attending a 12-week telecoaching program. Furthermore, the decline in adherence may be related to motivational problems. This is supported by the fact that the median adherence to exercise prescription in our study gradually decreased (as opposed to a sudden decrease because of symptoms). This gradual decrease in adherence is not unique to our study and was also found in other studies on home-based exercise programs in both CHF and COPD [25-27]. This highlights one of the most challenging aspects of CR and PR programs, which is to achieve actual behavioral changes that result in long-term maintenance of a certain activity level after completion of a rehabilitation program. This might be even more challenging in older patients with several comorbidities who have frequent relapses of their disease. Home-based rehabilitation programs should therefore be designed using effective behavior change techniques. Our study used the following behavior change techniques: goal setting, action planning, using graded tasks, self-monitoring of behavior, motivational interviewing during follow-up video calls, information about health consequences, and reviewing outcome goals. The number of techniques used is in line with other studies on eHealth physical activity interventions in patients with cardiovascular disease [28]. Behavior change techniques that might have been of added value in this study are the use of social support, providing prompts or cues to perform exercise, and providing rewards when an exercise session was completed (eg, by using text messages). Furthermore, a step-down approach to contact moments with the therapists and on-demand coaching may be more effective than a fixed coaching frequency and period. Nevertheless, the effectiveness of individual behavior change techniques with regard to physical activity outcome is not established yet [28].

Although digital technology can serve as a very useful medium to apply behavior change techniques, it should be noted that some of our study participants experienced difficulties using digital technology. Therefore, the technology can also serve as a barrier to achieving behavioral changes. In fact, Tadas et al [29] found that although the use of technology in eHealth interventions for patients with cardiovascular disease is readily accepted, the use of additional devices, especially in older adults, can be perceived as burdensome. Nevertheless, the facilitators with regard to the adoption of digital technology were background knowledge about the disease, in-the-moment understanding of health status by monitoring of physical (activity) data, the ability to connect and exercise with other patients and the involvement of family, receiving direct feedback, reminders through apps that might also use gamification principles, and the ability to personalize the program in terms of technical aspects [29]. The adoption of digital health technology in patients with COPD is associated with similar barriers and facilitators [30].

Program Completion and the COVID-19 Pandemic

The COVID-19 pandemic has added new barriers to participation in center-based CR and PR, namely temporary closure of CR and PR centers, reduced capacity after partial reopening of these centers, and anxiety of patients regarding regular visits to a CR or PR center [31]. This study was predominantly performed during the COVID-19 pandemic, but the inclusion of participants and follow-up consultations by the therapists were not affected by the pandemic. Moreover, adherence to exercise prescription is likely to be unaffected by the pandemic because exercise could be performed in the home environment, taking the patients' preferences into account. This study thereby shows the potential of home-based interventions as an alternative to center-based rehabilitation.

Program Satisfaction

Although the participants in our study were highly satisfied with the home-based exercise program, they were experiencing difficulties in particular with using the wrist-worn activity tracker, which was used to record their exercise sessions. To facilitate the activity tracker's use, one of its software features was adjusted before the study to enable the participants to start and end training sessions more easily. Moreover, functionalities that were not needed for study purposes were turned off. Nevertheless, some participants still experienced the use of the activity tracker as troublesome, which may have contributed to impaired adherence to the exercise prescription. Therefore, we suggest the use of sensors that patients only need to wear, without the need to perform additional actions. Moreover, automatic transfer of sensor data to a health platform should be provided. In addition, sensors that provide simple metrics (such as step count) and give direct feedback to users have been particularly graded as useful [24].

Program Efficacy

Program efficacy was assessed using both subjective and objective measures. Although the sample size was too small to evaluate program efficacy with sufficient statistical power, and a control group was lacking, we observed an improvement in

the PSC scores. Moreover, a trend toward improvement in the mMRC Dyspnea Scale and HRQoL scores was observed. No clinically relevant changes were observed in the objective measures (1-MSTS and handgrip strength tests). Although several studies on home-based CR and PR showed evidence of improvement in QoL and physical performance, as well as reduction in hospital admissions [32], this cannot yet be confirmed in our studied patient population, which is characterized by severe combined cardiopulmonary disease in combination with several other comorbidities and frequent exacerbations. This is supported by the study by Demeyer et al [33], which showed a significant improvement in physical activity level in patients with COPD across the whole spectrum during a telecoaching intervention but demonstrated attenuated effects in patients with a higher disease severity.

Future Research

This exploratory study aims to optimize the intervention for a larger trial that will assess the effectiveness of home-based rehabilitation in patients with severe combined cardiac and pulmonary disease. On the basis of the findings of this study, we suggest making the following adjustments to optimize the intervention:

- Integration of telemonitoring strategies and telerehabilitation, that is, creating the ability to adjust exercise prescription in line with fluctuations in symptoms and exercise tolerance.
- Using prompts to stimulate participants to perform their exercise and also offering rewards after completing an exercise session, for example, through text messages.
- Use of a sensor that is easy to use, does not need the participant to perform several additional actions, provides metrics that are easy to interpret, and offers direct feedback.
- A treatment period of 8 weeks was experienced as rather short by the therapists with regard to providing sufficient coaching and education for the participants to achieve their rehabilitation goals. Although a duration of 8 weeks is more often used in rehabilitation programs, a prolonged period of coaching (at least 12 weeks) in this population consisting of patients with complex health conditions might result in higher program efficacy. A step-down approach to contact moments with the therapists might be of added value to be able to extend the program duration without causing an additional financial burden. Moreover, it might stimulate participants to perform their exercise without frequent feedback from the therapists, thereby preparing them for the period after the rehabilitation program has ended.

Limitations

This study included a few limitations. First, the participants were already using the digital platform for telemonitoring purposes. The familiarity with the platform might have resulted in higher satisfaction rates with the home-based exercise program. Of the 10 participants, 1 (10%) was not familiar with the platform before beginning the program; however, this participant also scored high on satisfaction. Moreover, in a future study, home-based telerehabilitation will also be introduced after a period of telemonitoring because of the instability of the diseases.

Second, our study only focused on physical activity and not on other lifestyle domains. However, the ability to improve and maintain a certain activity level will also be influenced by other lifestyle behavior domains such as dietary behavior, mental stress, emotional well-being, and sleep quality. These factors were not taken into account in this rehabilitation program.

Finally, the results of this trial should be interpreted with caution because of the small number of participants and the lack of a control group. Nevertheless, because this study aims to explore program completion and adherence to, as well as patient satisfaction with, a home-based exercise program to facilitate

the design of a future randomized controlled trial, we chose not to include a control group.

Conclusions

Home-based telerehabilitation in older patients with combined severe cardiopulmonary disease and multiple other comorbidities is feasible in terms of a high program completion rate and high levels of patient satisfaction. However, a trend toward decline in adherence was observed during the program, which highlights the need for adding technology-based behavior change techniques in future trials.

Authors' Contributions

CH and HMCK contributed to the conception and design of the study. CH contributed to the acquisition of data. JGF and MS treated patients during the study. CH contributed to the analysis of data. All authors contributed to the interpretation of data. CH drafted the manuscript. All authors critically revised the manuscript. All gave final approval and agreed to be accountable for all aspects of this work, ensuring integrity and accuracy.

Conflicts of Interest

None declared.

Multimedia Appendix 1

An example of the Purpose-Designed Questionnaire.

[[DOCX File, 16 KB - formative_v5i11e28634_app1.docx](#)]

References

1. Roversi S, Fabbri LM, Sin DD, Hawkins NM, Agustí A. Chronic obstructive pulmonary disease and cardiac diseases. An urgent need for integrated care. *Am J Respir Crit Care Med* 2016 Dec 01;194(11):1319-1336. [doi: [10.1164/rccm.201604-0690SO](#)] [Medline: [27589227](#)]
2. Rabe KF, Hurst JR, Suissa S. Cardiovascular disease and COPD: dangerous liaisons? *Eur Respir Rev* 2018 Sep 30;27(149):180057 [FREE Full text] [doi: [10.1183/16000617.0057-2018](#)] [Medline: [30282634](#)]
3. Buddeke J, Bots ML, van Dis I, Visseren FL, Hollander M, Schellevis FG, et al. Comorbidity in patients with cardiovascular disease in primary care: a cohort study with routine healthcare data. *Br J Gen Pract* 2019 Jun;69(683):398-406 [FREE Full text] [doi: [10.3399/bjgp19X702725](#)] [Medline: [31064742](#)]
4. Ambrosetti M, Abreu A, Corrà U, Davos CH, Hansen D, Frederix I, et al. Secondary prevention through comprehensive cardiovascular rehabilitation: from knowledge to implementation. 2020 update. A position paper from the Secondary Prevention and Rehabilitation Section of the European Association of Preventive Cardiology. *Eur J Prev Cardiol* 2020 Mar 30;2047487320913379. [doi: [10.1177/2047487320913379](#)] [Medline: [32223332](#)]
5. Spruit MA, Singh SJ, Garvey C, ZuWallack R, Nici L, Rochester C, ATS/ERS Task Force on Pulmonary Rehabilitation. An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. *Am J Respir Crit Care Med* 2013 Oct 15;188(8):13-64. [doi: [10.1164/rccm.201309-1634ST](#)] [Medline: [24127811](#)]
6. de Araújo Pio C, Chaves GS, Davies P, Taylor RS, Grace SL. Interventions to promote patient utilisation of cardiac rehabilitation. *Cochrane Database Syst Rev* 2019 Feb 01;2:CD007131 [FREE Full text] [doi: [10.1002/14651858.CD007131.pub4](#)] [Medline: [30706942](#)]
7. van Engen-Verheul M, de Vries H, Kemps H, Kraaijenhagen R, de Keizer N, Peek N. Cardiac rehabilitation uptake and its determinants in the Netherlands. *Eur J Prev Cardiol* 2013 Apr;20(2):349-356. [doi: [10.1177/2047487312439497](#)] [Medline: [22345699](#)]
8. Ruano-Ravina A, Pena-Gil C, Abu-Assi E, Raposeiras S, van 't Hof A, Meindersma E, et al. Participation and adherence to cardiac rehabilitation programs. A systematic review. *Int J Cardiol* 2016 Nov 15;223:436-443. [doi: [10.1016/j.ijcard.2016.08.120](#)] [Medline: [27557484](#)]
9. Levack W, Watson J, Hay - Smith E, Davies C, Ingham T, Jones B, et al. Factors influencing referral to and uptake and attendance of pulmonary rehabilitation for chronic obstructive pulmonary disease: a qualitative evidence synthesis of the experiences of service users, their families, and healthcare providers. *Cochrane Database Syst Rev* 2018;11:CD013195. [doi: [10.1002/14651858.cd013195](#)]

10. Cox NS, Oliveira CC, Lahham A, Holland AE. Pulmonary rehabilitation referral and participation are commonly influenced by environment, knowledge, and beliefs about consequences: a systematic review using the Theoretical Domains Framework. *J Physiother* 2017 Apr;63(2):84-93 [FREE Full text] [doi: [10.1016/j.jphys.2017.02.002](https://doi.org/10.1016/j.jphys.2017.02.002)] [Medline: [28433238](https://pubmed.ncbi.nlm.nih.gov/28433238/)]
11. Anderson L, Sharp GA, Norton RJ, Dalal H, Dean SG, Jolly K, et al. Home-based versus centre-based cardiac rehabilitation. *Cochrane Database Syst Rev* 2017 Dec 30;6:CD007130. [doi: [10.1002/14651858.CD007130.pub4](https://doi.org/10.1002/14651858.CD007130.pub4)] [Medline: [28665511](https://pubmed.ncbi.nlm.nih.gov/28665511/)]
12. Cox NS, Dal Corso S, Hansen H, McDonald CF, Hill CJ, Zanaboni P, et al. Telerehabilitation for chronic respiratory disease. *Cochrane Database Syst Rev* 2021 Jan 29;1:CD013040. [doi: [10.1002/14651858.CD013040.pub2](https://doi.org/10.1002/14651858.CD013040.pub2)] [Medline: [33511633](https://pubmed.ncbi.nlm.nih.gov/33511633/)]
13. Reeves GR, Whellan DJ, O'Connor CM, Duncan P, Eggebeen JD, Morgan TM, et al. A novel rehabilitation intervention for older patients with acute decompensated heart failure: the REHAB-HF pilot study. *JACC Heart Fail* 2017 May;5(5):359-366 [FREE Full text] [doi: [10.1016/j.jchf.2016.12.019](https://doi.org/10.1016/j.jchf.2016.12.019)] [Medline: [28285121](https://pubmed.ncbi.nlm.nih.gov/28285121/)]
14. Bernocchi P, Vitacca M, La Rovere MT, Volterrani M, Galli T, Baratti D, et al. Home-based telerehabilitation in older patients with chronic obstructive pulmonary disease and heart failure: a randomised controlled trial. *Age Ageing* 2018 Dec 01;47(1):82-88. [doi: [10.1093/ageing/afx146](https://doi.org/10.1093/ageing/afx146)] [Medline: [28985325](https://pubmed.ncbi.nlm.nih.gov/28985325/)]
15. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982;5(3):233-237. [doi: [10.1016/0149-7189\(82\)90074-x](https://doi.org/10.1016/0149-7189(82)90074-x)] [Medline: [10259963](https://pubmed.ncbi.nlm.nih.gov/10259963/)]
16. de Brey H. A cross-national validation of the client satisfaction questionnaire: the Dutch experience. *Eval Program Plann* 1983;6(3-4):395-400. [doi: [10.1016/0149-7189\(83\)90018-6](https://doi.org/10.1016/0149-7189(83)90018-6)] [Medline: [10267266](https://pubmed.ncbi.nlm.nih.gov/10267266/)]
17. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest* 1988 Mar;93(3):580-586. [doi: [10.1378/chest.93.3.580](https://doi.org/10.1378/chest.93.3.580)] [Medline: [3342669](https://pubmed.ncbi.nlm.nih.gov/3342669/)]
18. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013 Sep;22(7):1717-1727 [FREE Full text] [doi: [10.1007/s11136-012-0322-4](https://doi.org/10.1007/s11136-012-0322-4)] [Medline: [23184421](https://pubmed.ncbi.nlm.nih.gov/23184421/)]
19. Bohannon RW, Crouch R. 1-Minute Sit-to-Stand Test: systematic review of procedures, performance, and clinometric properties. *J Cardiopulm Rehabil Prev* 2019 Jan;39(1):2-8. [doi: [10.1097/HCR.0000000000000336](https://doi.org/10.1097/HCR.0000000000000336)] [Medline: [30489442](https://pubmed.ncbi.nlm.nih.gov/30489442/)]
20. Roberts HC, Denison HJ, Martin HJ, Patel HP, Syddall H, Cooper C, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age Ageing* 2011 Jul;40(4):423-429 [FREE Full text] [doi: [10.1093/ageing/afr051](https://doi.org/10.1093/ageing/afr051)] [Medline: [21624928](https://pubmed.ncbi.nlm.nih.gov/21624928/)]
21. Zwisler A, Norton RJ, Dean SG, Dalal H, Tang LH, Wingham J, et al. Home-based cardiac rehabilitation for people with heart failure: a systematic review and meta-analysis. *Int J Cardiol* 2016 Oct 15;221:963-969. [doi: [10.1016/j.ijcard.2016.06.207](https://doi.org/10.1016/j.ijcard.2016.06.207)] [Medline: [27441476](https://pubmed.ncbi.nlm.nih.gov/27441476/)]
22. Tabak M, Brusse-Keizer M, van der Valk P, Hermens H, Vollenbroek-Hutten M. A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: a pilot randomized controlled trial. *Int J Chron Obstruct Pulmon Dis* 2014;9:935-944 [FREE Full text] [doi: [10.2147/COPD.S60179](https://doi.org/10.2147/COPD.S60179)] [Medline: [25246781](https://pubmed.ncbi.nlm.nih.gov/25246781/)]
23. Vasilopoulou M, Papaioannou AI, Kaltsakas G, Louvaris Z, Chynkiamis N, Spetsioti S, et al. Home-based maintenance tele-rehabilitation reduces the risk for acute exacerbations of COPD, hospitalisations and emergency department visits. *Eur Respir J* 2017 May;49(5):1602129 [FREE Full text] [doi: [10.1183/13993003.02129-2016](https://doi.org/10.1183/13993003.02129-2016)] [Medline: [28546268](https://pubmed.ncbi.nlm.nih.gov/28546268/)]
24. Loeckx M, Rabinovich RA, Demeyer H, Louvaris Z, Tanner R, Rubio N, et al. Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: mixed-methods study on patient experiences and lessons for implementation. *JMIR Mhealth Uhealth* 2018 Dec 21;6(12):e200 [FREE Full text] [doi: [10.2196/mhealth.9774](https://doi.org/10.2196/mhealth.9774)] [Medline: [30578215](https://pubmed.ncbi.nlm.nih.gov/30578215/)]
25. Jolly K, Taylor RS, Lip GY, Davies M, Davis R, Mant J, et al. A randomized trial of the addition of home-based exercise to specialist heart failure nurse care: the Birmingham Rehabilitation Uptake Maximisation study for patients with Congestive Heart Failure (BRUM-CHF) study. *Eur J Heart Fail* 2009 Feb;11(2):205-213 [FREE Full text] [doi: [10.1093/eurjhf/hfn029](https://doi.org/10.1093/eurjhf/hfn029)] [Medline: [19168520](https://pubmed.ncbi.nlm.nih.gov/19168520/)]
26. O'Connor CM, Whellan DJ, Lee KL, Keteyian SJ, Cooper LS, Ellis SJ, HF-ACTION Investigators. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. *J Am Med Assoc* 2009 Apr 08;301(14):1439-1450 [FREE Full text] [doi: [10.1001/jama.2009.454](https://doi.org/10.1001/jama.2009.454)] [Medline: [19351941](https://pubmed.ncbi.nlm.nih.gov/19351941/)]
27. Bourne S, DeVos R, North M, Chauhan A, Green B, Brown T, et al. Online versus face-to-face pulmonary rehabilitation for patients with chronic obstructive pulmonary disease: randomised controlled trial. *BMJ Open* 2017 Jul 17;7(7):e014580 [FREE Full text] [doi: [10.1136/bmjopen-2016-014580](https://doi.org/10.1136/bmjopen-2016-014580)] [Medline: [28716786](https://pubmed.ncbi.nlm.nih.gov/28716786/)]
28. 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/)]
29. Tadas S, Coyle D. Barriers to and facilitators of technology in cardiac rehabilitation and self-management: systematic qualitative grounded theory review. *J Med Internet Res* 2020 Nov 11;22(11):e18025 [FREE Full text] [doi: [10.2196/18025](https://doi.org/10.2196/18025)] [Medline: [33174847](https://pubmed.ncbi.nlm.nih.gov/33174847/)]

30. Slevin P, Kessie T, Cullen J, Butler MW, Donnelly SC, Caulfield B. A qualitative study of chronic obstructive pulmonary disease patient perceptions of the barriers and facilitators to adopting digital health technology. *Digit Health* 2019;5:2055207619871729 [FREE Full text] [doi: [10.1177/2055207619871729](https://doi.org/10.1177/2055207619871729)] [Medline: [31489206](https://pubmed.ncbi.nlm.nih.gov/31489206/)]
31. Scherrenberg M, Wilhelm M, Hansen D, Völler H, Cornelissen V, Frederix I, et al. The future is now: a call for action for cardiac telerehabilitation in the COVID-19 pandemic from the secondary prevention and rehabilitation section of the European Association of Preventive Cardiology. *Eur J Prev Cardiol* 2020 Jul 02:2047487320939671. [doi: [10.1177/2047487320939671](https://doi.org/10.1177/2047487320939671)] [Medline: [32615796](https://pubmed.ncbi.nlm.nih.gov/32615796/)]
32. Aragaki D, Luo J, Weiner E, Zhang G, Darvish B. Cardiopulmonary telerehabilitation. *Phys Med Rehabil Clin N Am* 2021 May;32(2):263-276 [FREE Full text] [doi: [10.1016/j.pmr.2021.01.004](https://doi.org/10.1016/j.pmr.2021.01.004)] [Medline: [33814057](https://pubmed.ncbi.nlm.nih.gov/33814057/)]
33. Demeyer H, Louvaris Z, Frei A, Rabinovich RA, de Jong C, Gimeno-Santos E, Mr Papp PROactive Study Group and the PROactive Consortium. Physical activity is increased by a 12-week semiautomated telecoaching programme in patients with COPD: a multicentre randomised controlled trial. *Thorax* 2017 May;72(5):415-423 [FREE Full text] [doi: [10.1136/thoraxjnl-2016-209026](https://doi.org/10.1136/thoraxjnl-2016-209026)] [Medline: [28137918](https://pubmed.ncbi.nlm.nih.gov/28137918/)]

Abbreviations

1-MSTS: 1-minute sit-to-stand
CHF: chronic heart failure
COPD: chronic obstructive pulmonary disease
CR: cardiac rehabilitation
CSQ: Client Satisfaction Questionnaire
EQ-5D-5L: EuroQol 5-Dimension, 5-Level
HRQoL: health-related quality of life
mMRC: modified Medical Research Council
PDQ: Purpose-Designed Questionnaire
PR: pulmonary rehabilitation
PSC: Patient-Specific Complaints

Edited by G Eysenbach; submitted 15.03.21; peer-reviewed by Z Louvaris, M Scherrenberg, S Tadas; comments to author 27.04.21; revised version received 01.06.21; accepted 16.06.21; published 09.11.21.

Please cite as:

Herkert C, Graat-Verboom L, Gilsing-Fernhout J, Schols M, Kemps HMC

Home-Based Exercise Program for Patients With Combined Advanced Chronic Cardiac and Pulmonary Diseases: Exploratory Study
JMIR Form Res 2021;5(11):e28634

URL: <https://formative.jmir.org/2021/11/e28634>

doi: [10.2196/28634](https://doi.org/10.2196/28634)

PMID: [34751655](https://pubmed.ncbi.nlm.nih.gov/34751655/)

©Cyrille Herkert, Lidwien Graat-Verboom, Judith Gilsing-Fernhout, Manon Schols, Hareld Marijn Clemens Kemps. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 09.11.2021. 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 Text Message Intervention for Adolescents With Depression and Their Parents or Caregivers to Overcome Cognitive Barriers to Mental Health Treatment Initiation: Focus Groups and Pilot Trial

Brian Suffoletto¹, MS, MD; Tina Goldstein², PhD; David Brent³, MD

¹Department of Emergency Medicine, Stanford University, Palo Alto, CA, United States

²Department of Psychology, University of Pittsburgh, Pittsburgh, PA, United States

³Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:

Brian Suffoletto, MS, MD

Department of Emergency Medicine

Stanford University

900 Welch Road, Suite 350

Palo Alto, CA, 94304

United States

Phone: 1 4129016892

Email: suffbp@stanford.edu

Abstract

Background: Many adolescents with depression do not pursue mental health treatment following a health care provider referral. We developed a theory-based automated SMS text message intervention (Text to Connect [T2C]) that attempts to reduce cognitive barriers to the initiation of mental health care.

Objective: In this two-phase study, we seek to first understand the potential of T2C and then test its engagement, usability, and potential efficacy among adolescents with depression and their parents or caregivers.

Methods: In phase 1, we conducted focus groups with adolescents with depression (n=9) and their parents or caregivers (n=9) separately, and transcripts were examined to determine themes. In phase 2, we conducted an open trial of T2C comprising adolescents with depression referred to mental health care (n=43) and their parents or caregivers (n=28). We assessed usability by examining program engagement, usability ratings, and qualitative feedback at the 4-week follow-up. We also assessed potential effectiveness by examining changes in perceived barriers to treatment and mental health care initiation from baseline to 4 weeks.

Results: In phase 1, we found that the themes supported the T2C approach. In phase 2, we observed high engagement with daily negative affect check-ins, high usability ratings, and decreased self-reported barriers to mental health treatment over time among adolescents. Overall, 52% (22/42) of the adolescents who completed follow-up reported that they had attended an appointment with a mental health care specialist. Of the 20 adolescents who had not attended a mental health care appointment, 5% (1/20) reported that it was scheduled for a future date, 10% (2/20) reported that the primary care site did not have the ability to help them schedule a mental health care appointment, and 15% (3/20) reported that they were no longer interested in receiving mental health care.

Conclusions: The findings from this study suggest that T2C is acceptable to adolescents with depression and most parents or caregivers; it is used at high rates; and it may be helpful to reduce cognitive barriers to mental health care initiation.

(*JMIR Form Res* 2021;5(11):e30580) doi:[10.2196/30580](https://doi.org/10.2196/30580)

KEYWORDS

adolescent; depression; help seeking; text message; intervention

Introduction

Background

Approximately 6% of US adolescents aged between 12 and 17 years have a diagnosis of depression; however, >20% of those diagnosed with depression have not received treatment in the past year [1]. Furthermore, studies show that, on average, 10 years elapse from the time of mental health symptom onset to receipt of treatment [2]. Among those referred to mental health care for depression or suicidality, only 18% access mental health care within 6 months [3], which may increase only to 32% with integrated mental health services [4]. If left untreated, depression can become harder to manage [5] and puts individuals at greater risk for suicide [6].

There are many documented reasons why adolescents do not seek or receive mental health treatment. Help-seeking for mental health symptoms is a multistep process beginning with an awareness of the problem, followed by an expression of the problem and the need for help to others—which for adolescents is typically a parent or caregiver—followed by a discussion and plan for help-seeking [7]. Barriers to advancement along this process for mental health treatment include lack of health literacy, perceived stigma, concerns about confidentiality, and preference for self-reliance. These barriers are amplified in the adolescent development period [8] and are difficult to overcome when both the child and parent or caregiver may have different beliefs or barriers.

Interventions aimed at improving help-seeking for mental health have shown mixed results. A systematic review from 2012 demonstrated that mental health help-seeking interventions improve attitudes but not necessarily behaviors among young people [9]. A more recent systematic review and meta-analysis of 98 mental health help-seeking interventions found that formal help-seeking behaviors improved when the intervention was delivered to adults with or at risk of mental health problems but not when delivered to adolescents [10]. Among a review of interventions targeting mental health in adolescents specifically, almost all were delivered in schools, and three-quarters were education-based [11].

Digital behavioral interventions can overcome barriers of in-person or location-based interventions [12]; however, they have also shown mixed promise for adolescent help-seeking for mental health. In 2014, a systematic review of 18 studies of web-based interventions targeting help-seeking in young people aged between 14 and 25 years (only 4 of which included adolescents) found that although the trials did not change in help-seeking behavior, quasi-experimental and cross-sectional studies reported that interventions facilitated seeking professional help for an average of 35% of users [12]. A recent scoping review of 4 studies of digital health interventions aimed at improving help-seeking behavior or access to mental health services among parents or caregivers of 2- to 12-year-olds concluded there was some evidence of improved mental health literacy but not necessarily help-seeking [13].

Objectives

Given the current gap in evidence for help-seeking interventions for mental health care initiation among at-risk adolescents, we developed and tested an automated SMS text message intervention (Text To Connect [T2C]) designed to reduce cognitive barriers to mental health care help-seeking among adolescents with depression and their parents or caregivers. This report details the design of our intervention and presents the results from a two-phase study. In phase 1, we conducted focus groups with adolescents with depression and their parents or caregivers separately to understand the barriers to mental health treatment and the acceptability of the T2C approach. In phase 2, we conducted an open trial of T2C for adolescents with depression referred to mental health care and their parents or caregivers. We assessed usability by examining T2C program engagement as well as usability ratings and qualitative feedback at the 4-week follow-up. We assessed potential effectiveness by examining changes in perceived barriers to mental health care and mental health care initiation from baseline to 4-weeks. Findings from this study are critical to inform the further development and evaluation of digital interventions aimed at improving help-seeking and linkage to mental health care.

Methods

T2C Intervention Design

Theoretical Basis

We designed T2C to reduce cognitive barriers to seeking mental health care. T2C was informed by the Health Belief Model [14], where perceived susceptibility or severity of depression, perceived benefits to mental health treatment, and cues to action all influence the likelihood of mental health care help-seeking. The T2C design was also shaped by the belief that help-seeking is a multistep process that starts with awareness of the problem, followed by a recognition that mental health care can help and subsequently by a discussion and plan for help-seeking between child and parents or caregivers [7].

Technical Design

T2C is an automated computer program created and housed in the Office of Academic Computing at the University of Pittsburgh Department of Psychiatry. In brief, it is a table-driven program that uses Microsoft SQL, which batches outgoing timed messages and receives incoming messages using Twilio (Twilio Inc). Program flow, branching logic, query wording, and feedback libraries were all created before the enrollment of the first participant. Although the program was designed to function with the child-parent or -caregiver dyad, we allowed either parent or caregiver or child to enroll without the other to enhance applicability to real-world situations.

Onboarding

To onboard, T2C participants were prompted to text a unique keyword to initiate T2C. Once initiated, participants received introductory texts including instructions that they could drop out of the T2C program at any time by texting *Quit*. Parents or caregivers and adolescents received a series of SMS text

message queries to assist in the stratification and tailoring of T2C content.

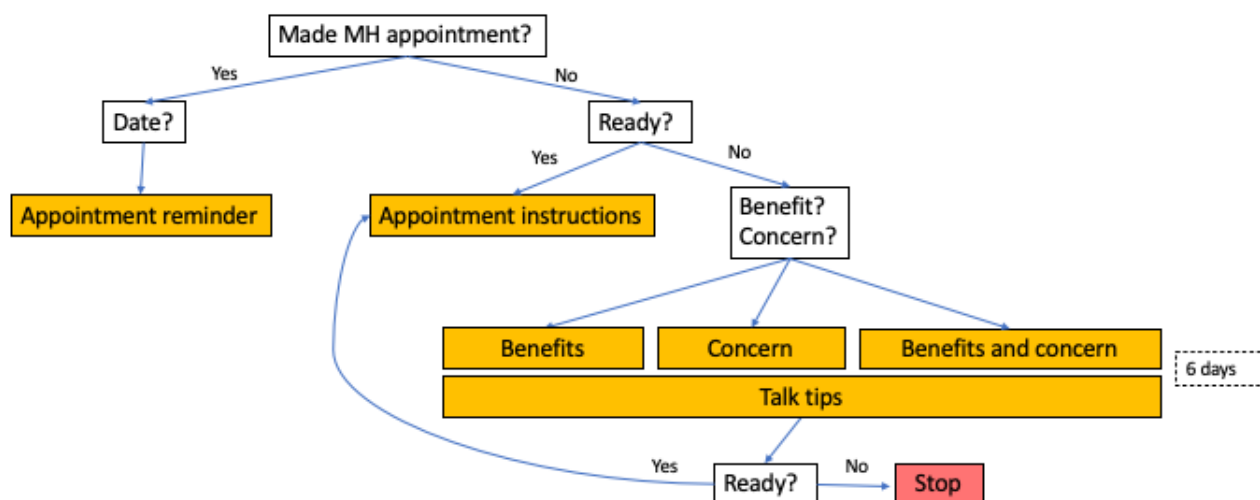
T2C Intervention

Parent or Caregiver

Upon enrollment (Figure 1), parents or caregivers received the message, “Have you made an appointment for your child to see one of our mental health clinicians?” If they reported that they had already made an appointment, they were sent a link to an informational page about the mental health clinicians at their site and a query, “What is the date of your appointment?” which was used to send a reminder text the day before the scheduled appointment. If the parent or caregiver reported not having made an appointment, they were asked, “Are you ready to make an appointment for your child?” If they replied affirmatively, they were provided instructions on how to set up a mental health care appointment. If they indicated not being ready, they received the messages, “Do you feel that your child could benefit

from seeing a mental health clinician?” and “Do you have concerns about your child seeing a mental health clinician?” On the basis of their responses to these queries, an individual would receive psychoeducation aimed at reducing stigma, increasing perceived benefits of mental health care, or both. Psychoeducation was designed in a microlearning format [15] as daily *true* or *false* questions for 6 days (see Multimedia Appendix 1 for sample messages). To encourage conversations between parent or caregiver and child regarding depression and mental health care, T2C provided cues to action at the end of each day in the form of a *talk tip*, for example, “Sometimes starting the conversation is hard. Consider: Are you okay? I’m here if you want to talk.” On day 7, parent or caregiver participants received a check-in asking if they were then ready to make an appointment. At this point, if they did not feel ready, they were referred to their child’s primary care physician to discuss their decisions. If they indicated that they were ready, they received appointment instructions.

Figure 1. Parent or caregiver Text To Connect flow diagram. MH: mental health.

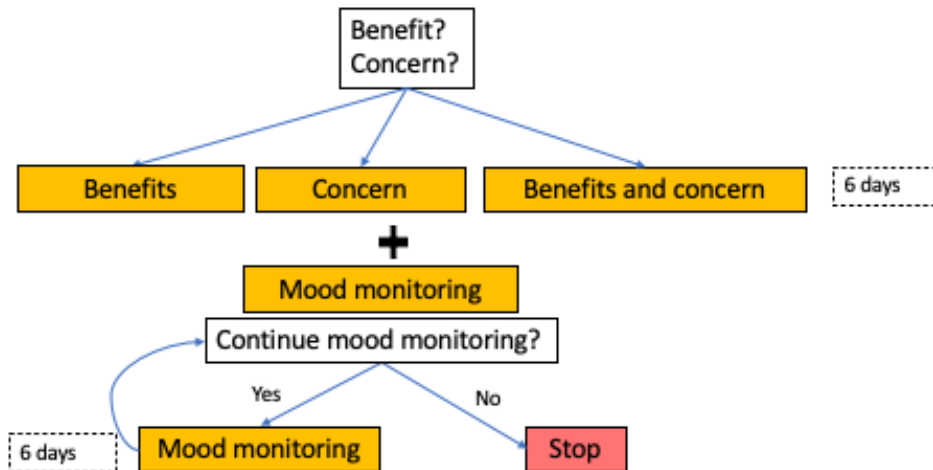


Adolescent

Upon enrollment (Figure 2), adolescent participants received similar introductory queries to their parents or caregivers to help tailor intervention material: “Do you think that seeing a mental health clinician could help you?” and “Do you have any concerns about seeing a mental health clinician?” On the basis of their responses to these queries, an adolescent would receive psychoeducation in the form of *true* or *false* questions for 6 days aimed at reducing stigma, increasing perceived benefits of mental health care, or both (see Multimedia Appendix 1 for sample messages). To increase awareness of symptoms and severity of depression among adolescents, T2C incorporated daily mood monitoring, which has been shown to differentiate adolescents at higher risk for psychopathology [16]; “How often have you felt down, depressed or sad today on a scale 1 (not at all) to 5 (all day)?” As depressed adolescents often

underestimate the severity of their symptoms and do not correctly perceive their degree of psychological risk associated with depressive symptoms [17], T2C provided adolescents with both immediate and end-of-week feedback on their depression-related symptoms. Immediate feedback focused on relating symptoms to underlying psychopathology, for example, “You report some symptoms of depression today” and prompting self-management strategies, for example, “Try to focus on one small thing that was good today.” End-of-week feedback focused on quantifying the extent of depressive symptoms; “You reported some symptoms of depression on [X%] of days this past week,” where X was calculated as: (days with any depressive symptoms) / (days with any report). Adolescents were then offered the option of continuing mood monitoring: “Are you interested in continuing to monitor your mood for another week?”

Figure 2. Adolescent Text To Connect flow diagram.



Safety

We included a safety feature whereby any response from adolescents or parents or caregivers that included a concerning keyword (ie, help, suicide, kill, hurt, die, pills, gun, and please) immediately triggered an automated message providing crisis resources: “If you are experiencing an emergency, please call 9-1-1 or text CONNECT to 741741 to reach the Crisis Text Line.”

Phase 1: Focus Groups

Study Design

Phase 1 was a focus group study of adolescent-parent or caregiver dyads. All participants completed written informed consent. All procedures were approved by the institutional review board of the University of Pittsburgh.

Participants

We recruited adolescent-parent or caregiver dyads from primary care clinics and through flyers posted in clinical settings at a single health system in western Pennsylvania. Research assistants approached families or contacted them by phone after they expressed interest in participating in the research. Eligible adolescent-parent or caregiver dyads were included if the child was aged between 12 and 17 years and screened positive for depressive symptoms as part of standard clinical care at their pediatric primary care office. Participants were not eligible if they were not fluent in English. Participants were scheduled for a focus group session where signed consent was obtained from young people aged >16 years and signed parent or caregiver consent and child assent from those aged <16 years.

Procedures

Before conducting any focus groups, a standardized, semistructured qualitative guide was developed to increase consistency across the interview sessions. Data collection for adolescents and parents or caregivers occurred between November 2018 and January 2019. The focus groups took place at research staff offices, and the sessions lasted approximately 120 minutes. Adolescent focus group sessions occurred separately from the parent or caregiver sessions. The focus groups were conducted by a female facilitator with expertise in

focus group techniques and qualitative methodology. All sessions were audio recorded digitally and transcribed. A slide presentation of a T2C mockup was shown to the participants as part of the session. The interview topic guide covered (1) barriers to mental health care, (2) ways to motivate teens to initiate mental health care, (3) reactions to the T2C prototype, and (4) design areas to develop further regarding the design and content of T2C. All participants received US \$25 for focus group participation.

Analyses

Focus group transcripts were analyzed using a thematic content analysis approach and used the qualitative research software package ATLAS.ti 5.0 (Scientific Software Development). A preliminary codebook was created based on close readings of the first transcripts, incorporating explicit domains from interview guides (deductive themes) as well as recurrent unanticipated themes that emerged across transcripts (inductive themes). The coded text was further reviewed through an iterative process, resulting in refined themes. We did not record which individual participant said which statement or counted how many participants agreed or disagreed with a given statement. In presenting the results, we chose participant quotes that represented both most sentiments within each theme as well as any quote that offered a contrasting opinion within that theme.

Phase 2: Pilot Single-Arm Trial

Study Design

We then conducted an open single-arm trial of T2C, preregistered at clinicaltrials.gov (NCT04560075). Data were collected at baseline and 4-weeks post enrollment. Adolescents could enroll without their parents or caregivers enrolling. Participants completed written informed consent. All procedures were approved by the institutional review board of the University of Pittsburgh.

Participants

We recruited adolescents from 4 primary care sites (22/43, 51%), 1 specialty clinic (2/43, 5%), and 1 mental health clinic (19/43, 44%) in Pittsburgh, PA. Participating sites were asked to identify and refer adolescents with depression who were referred to mental health care according to their usual clinical

practice. Upon identifying potential adolescent participants, the care providers asked if the parent or caregiver was willing to speak with a researcher to learn more about the study and provide assent for their child. Once an adolescent was enrolled, an available and interested parent or caregiver (one per child) was assessed for inclusion and, if verified, were enrolled. Once informed consent was obtained, the research team scheduled baseline assessment and relevant onboarding.

Inclusion criteria were as follows: (1) adolescents aged 12-17 years, (2) English language fluency and literacy sufficient to engage in study assessment and intervention, (3) adolescent Patient Health Questionnaire-9 > 11 or Patient Health Questionnaire-9 item 9 > 1, and (4) adolescent referred for mental health services by their primary care provider or other health care provider. Exclusion criteria were as follows: adolescents and parents or caregivers who do not own a cell phone with SMS text message capabilities.

Measures

Cognitive barrier classification (ie, perceived benefits or stigma) and program engagement data (ie, negative affect check-ins) were obtained from SMS text messaging logs. Demographics, program usability, barriers to mental health care, and health care use measures were obtained through a secure website at baseline and 4-weeks from adolescents.

T2C Usability

T2C usability was measured with questions taken from the Post-Study System Usability Scale [18]. The questions included “Overall, I am satisfied with how easy it is to use Text2Connect,” “The information provided with Text2Connect was clear,” and “I liked interacting with the Text2Connect Program,” all with response options ranging from 1=strongly disagree to 7=strongly agree. The question “I needed to learn a lot of things before I could get going with Text2Connect” was presented with response options ranging from 1=strongly disagree to 5=strongly agree. The question “If a friend were in need of a mental health care, would you recommend Text2Connect to help him or her manage symptoms?” was presented with response options ranging from 1=no, definitely not; 2=no, I don’t think so; 3=yes, I think so; and 4=yes, definitely, reduced in analyses to a dichotomous variable of yes or no.

Barriers to Mental Health Care

Barriers to mental health care were measured using a brief version of the Barriers to Adolescents Seeking Help scale, derived from the longer scale developed by Kuhl et al [19]. The abbreviated measure comprised 5 of the original 37 self-report items and specifically targets belief-based barriers to seeking professional mental health help (eg, “A therapist might make me do or say something that I don’t want to” and “If I had a problem and told a therapist, they would not keep it a secret”). Each item was presented with response options ranging from 1=strongly disagree to 6=strongly agree. Higher scores indicate higher belief-based barriers to professional mental health help-seeking. In this study, Cronbach α for the 5-item abbreviated measure Barriers to Adolescents Seeking Help was .66 (for adolescents) and .45 (for parents or caregivers). Because

internal consistency was low, all item responses are presented separately, and, for simplicity of presentation, we dichotomized responses into agree (if responded agree or strongly agree) and disagree (if responded slightly disagree, disagree, or disagree strongly).

Analysis

We assessed engagement by calculating the response rates to SMS text messages within and between individuals. We assessed usability and perceived utility by examining ratings and qualitative feedback from adolescents and parents or caregivers at the 4-week follow-up. Potential effectiveness was assessed by examining changes in perceived barriers to mental health care in adolescents and parents or caregivers from baseline to 4-weeks. We explored the patient characteristics associated with mental health care initiation at 4-weeks using chi-square tests.

Results

Phase 1: Focus Groups

Sample Characteristics

A total of 9 dyads of adolescents with depression and their parents or caregivers completed a focus group session. Adolescents were aged between 15 and 17 years; 33% (3/9) were female, 33% (3/9) were Black, and 22% (2/9) were more than one race. Parent or caregiver age ranged from 31 to 61 years; 89% (8/11) were female, 22% (2/9) were Black, and 9% (1/11) were more than one race.

Theme 1: Parents or Caregivers and Adolescents Face Multiple Barriers to Receiving Mental Health Care

Parents described playing an active role in scheduling appointments but experienced logistical barriers to successfully navigating the system. A mother stated the following:

You know that’s already like kind of an emotionally overwhelming sort of thing to be facing. Um, and to have like all the logistics piled on top with like it’s so complex and you have to navigate it all by yourself.

Parents described a dearth of mental health professionals specializing in adolescents, a need to find the *right fit* with their child’s personality, and a sense of frustration with the need for multiple mental health care providers. For example, one parent quipped, “One for evaluation, one for talk therapy, one for medication.” Parents and adolescents lamented the difficulty in fitting mental health care into their schedules. One parent even described the frustration with their child being *fired* when they missed an mental health care appointment.

Theme 2: Most Parents Report Being Supportive of Mental Health Care, but Adolescents Were Frequently Opposed to or Ambivalent About Seeking Mental Health Care

Common reasons adolescents reported being opposed to mental health care included the following: not believing that mental health treatment was necessary, not believing that mental health treatment is effective, not knowing what to expect from therapy, fearing that the counselor would not respect their privacy, and

the concern that getting treatment means something is *wrong* with them. An adolescent bluntly summed this up as follows:

Counseling never helps, it's useless, it's a waste of time. I do not wanna go talk to somebody. Um, they do not know anything.

Theme 3: Adolescents Suggested Many Ways Mental Health Care Could Be Viewed as More Useful

One suggestion was to explain to an adolescent exactly what mental health care actually entailed. As one teen put it, a program should help them realize that through therapy, “they’ll be a better version of themselves [who] will be more aware of how to deal with their emotions properly rather than doing something they regret.” Another suggestion was to present them with evidence of their depression or negative mood to make treatment seem more necessary. A third suggestion was to help the teen find a therapist with whom they could build rapport. Finally, there was the idea that if all else failed, to incentivize the teen through payment. A parent echoed this idea when she said the following:

I think I used bribery, [Laughter] threats of a privilege loss. Um, I'm trying to remember exactly what I did but I-I can't um, I mean I told him that he needed to go, like he just, he needed to uh, he needed to give it a shot, go at least once or twice.

Theme 4: Most Participants Had a Positive Reaction to T2C, and Many Thought That Using T2C Would Achieve the Goal of Increasing the Number of Teens Who Received Treatment

Texting was perceived as an ideal modality as it is easy, relatable, and simpler than talking on the phone. An adolescent said the following:

I think this is fabulous, um every teenager I see nowadays is on their phone non-stop, and I think that they can relate to some—cuz it's not a live person too. They're a little bit more—it's not a parent, it's

not a live person, so if they are having a feeling, they can say “yes” and not like that fear of uh the adult either standing there or like a parent like, you know. Yeah, so I actually love this a lot.

Parents thought it could be useful to convince reluctant parents that their children needed treatment and would provide parents with the same information as their children so that they could be *on the same page*, and also so that the parents could discuss the information with the child. There seemed to be mixed feelings about interacting with an automated communication system or *bot*; they expressed that they were more likely to disclose *true* feelings but also less likely to feel obligated to engage.

Theme 5: Adolescents and Parents Had Several Suggestions on How to Improve T2C

Adolescents felt that it was important to be very clear in onboarding about who would see their texted responses, especially responses related to their mood reports. Several adolescents suggested that messages may sometimes not grab their attention and could get dull over time and recommended using humor and emojis to get and keep the teens’ interest. Most adolescents did not like the *true* or *false* format of delivering psychoeducation, relating that some teens may get upset in being told they were *wrong*. Contrary to adolescent concerns about keeping their reports private, several parents thought that there would be value in having access to their child’s mood reports. A parent stated that shared reports could convince “reluctant parents that their children needed treatment...so that they could be on the same page, and also so that the parents could discuss the information with the child.”

Phase 2: Pilot Single-Arm Trial

Participants

Adolescents

Of the total sample of consented adolescent participants (n=43), most were born female (35/43, 81%) and White (33/43, 77%; [Table 1](#)).

Table 1. Characteristics of phase-2 study participants (N=71).

Characteristics	Adolescent (n=43)	Parent or caregiver (n=28)
Age (years), mean (SD)	16 (2.9)	— ^a
Sex at birth (female), n (%)	35 (81)	25 (89)
Race, n (%)		
White	33 (77)	27 (96)
Black or African American	3 (7)	1 (4)
Asian	2 (5)	0 (0)
More than one	2 (5)	0 (0)
Other	3 (7)	0 (0)
Hispanic	2 (5)	NR
Baseline depressive symptoms (PHQ-9 ^b), mean (SD)	8.9 (5.5)	NR
Lifetime suicide attempts, mean (SD)	2 (1.5)	NR
Household income (US \$), n (%)		
<25,000	NR	3 (11)
25,000-50,000	NR	5 (18)
50,000-75,000	NR	3 (11)
75,000-100,000	NR	5 (18)
100,000	NR	12 (43)

^aNot available.

^bPHQ-9: Patient Health Questionnaire-9 item.

Parents or Caregivers

A total of 28 parents or caregivers agreed to participate (18/28, 65% of dyads). Parents or caregivers were mostly female (25/28, 89%) and White (27/28, 96%). Approximately 43% (12/28) reported living in a household that makes >US \$100,000 per year.

Mental Health Care and Barrier Classification

Adolescents

At baseline, before any intervention exposure, 21% (9/43) of adolescents reported not perceiving benefits to mental health care and 21% (9/43) reported perceived stigma to mental health care.

Parent or Caregivers

At baseline, before any intervention exposure, 68% (19/28) of parents or caregivers reported already making a first appointment for mental health care for their child. Of the 9 remaining parents or caregivers, 5 (56%) were ready to make

an appointment. Of the 4 remaining parents or caregivers who did not feel ready, 1 (25%) reported not perceiving benefits to mental health care for their child, and 1 (25%) perceived stigma to mental health care. Among the 4 matched adolescents whose parents or caregivers were not ready, 2 (50%) adolescents reported not perceiving benefits to mental health care and 1 (25%) reported perceived stigma to mental health care.

Adolescent Negative Affect Check-ins

A total of 1848 SMS text message negative affect check-ins were sent over the study period and 89.23% (1649/1848) were completed. After completing the first week of T2C engagement, 79% (34/43) of adolescents opted to continue mood monitoring. The median number of weeks an adolescent opted to continue the T2C intervention was 4 weeks and a maximum of 35 weeks in 1 participant. The distribution of responses by adolescent participants to the daily negative affect check-in “How often have you felt down, depressed or sad today on a scale 1 (not at all) to 5 (all day)?” across the first 4 weeks are shown in [Figure 3](#).

Figure 3. Daily negative affect reports by participant.

Program Usability

Overall, 93% (40/43) of adolescents and 96% (27/28) of parents or caregivers completed the usability survey questions (taken

from the Post-Study System Usability Scale) at the 4-week follow-up. Overall, mean usability scores were high among adolescents and parents (Table 2).

Table 2. Program usability (N=67).

Usability survey questions	Adolescent (n=40)	Parent or caregiver (n=27)
Overall, I am satisfied with how easy it is to use Text2Connect, mean (SD) ^a	6.1 (1.3)	6.5 (1.0)
The information provided with Text2Connect was clear, mean (SD) ^a	6.3 (0.9)	6.3 (1.3)
I liked interacting with the Text2Connect Program, mean (SD) ^a	5.1 (1.5)	5.22 (1.9)
If a friend were in need of a mental health care, would you recommend Text2Connect to help him or her manage symptoms? n (%) ^b	28 (70)	21 (78)
I needed to learn a lot of things before I could get going with Text2Connect, mean (SD) ^c	1.5 (1.0)	1.1 (0.3)

^aRatings for statements 1 through 3: 1=strongly disagree to 7=strongly agree.

^bRatings for statement 4: 1=no, definitely not; 2=no, I don't think so; 3=yes, I think so; and 4=yes, definitely was reduced to yes or no and presented as n (%) responding yes.

^cRatings for statement 5: 1=strongly disagree to 5=strongly agree.

Qualitative Feedback on T2C

Adolescents

When asked why they would recommend or not recommend the T2C program, adolescents expressed a range of opinions and thoughts. Among those who would recommend the program, several highlighted the benefits of check-ins, noting it made them feel "less alone" and "cared for." Several others noted the ease of interacting with the program. Among adolescents who

would not recommend the T2C program, 1 participant stated, "I think it's good for tracking your mood and how you're feeling on a regular daily basis but it's not actually helping you cope with your feelings." Another stated, "I would say it you're really in a bad mental state having an automated bot is not the best thing to have because it doesn't feel special to you, no real human connection."

Parent or Caregivers

Among those who would recommend the program, parents or caregivers noted the ease of use, 1 parent stating, “It was more laid back for my son than talking to someone in person.” Another commented, “It made you think about the scenarios. easier to open up about things. made recommendations. Text is easy these days and how kids like to communicate.” Several of those who would not recommend the T2C program commented that automation was limiting; one parent or caregiver stated as follows:

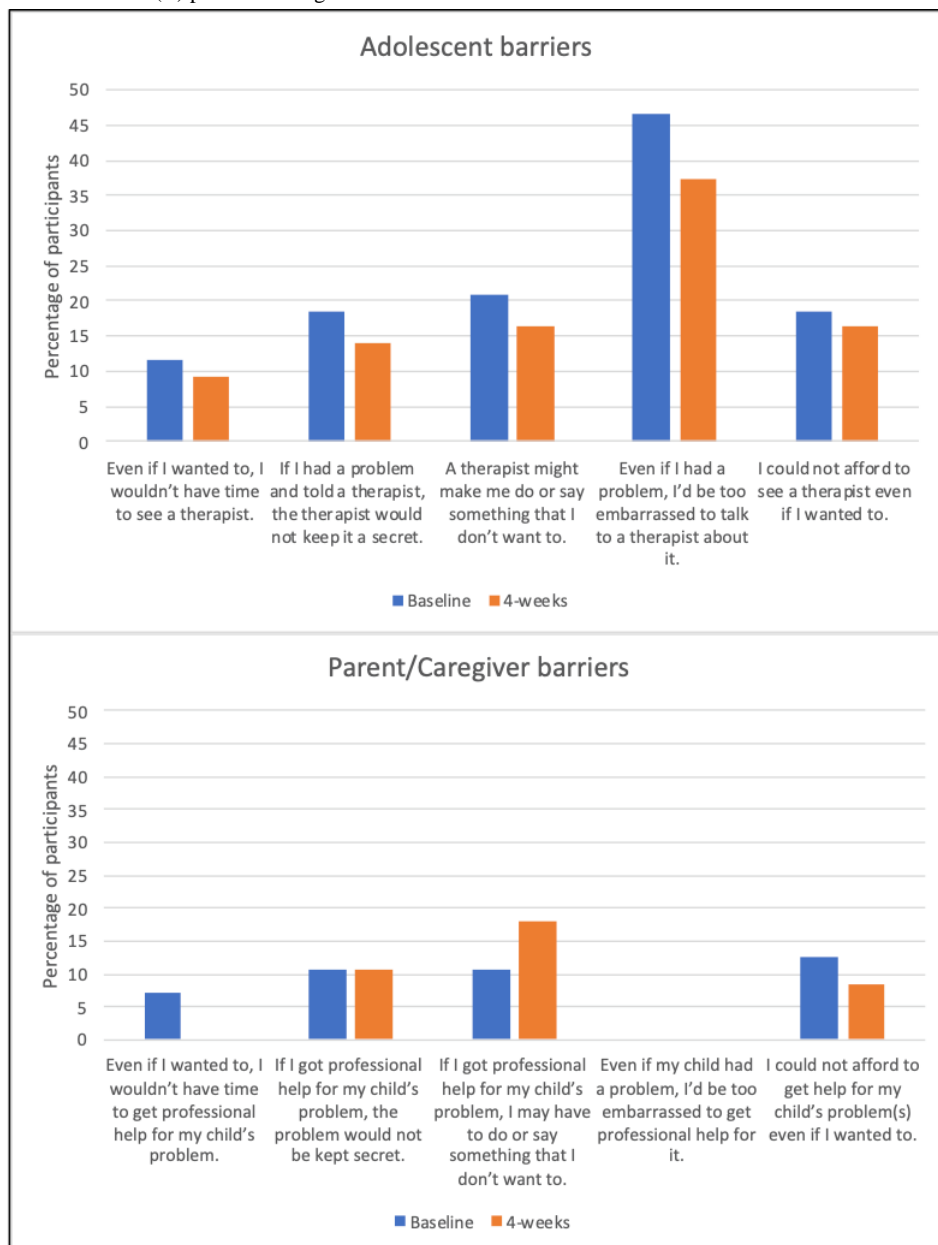
Not really interactive. Try to respond to text...we don't understand response call 911 if emergency. Disjointed. If someone was reaching out for help it wouldn't be helpful.

Barriers to Mental Health Care

Adolescents

At baseline (Figure 4), embarrassment was the most common barrier to mental health care, and lack of time was the least common barrier. There were small reductions (2%-10%) in all barriers to mental health care between baseline and the 4-week follow-up.

Figure 4. (A) Adolescent barriers and (B) parent or caregiver barriers.



Parent or Caregivers

At baseline, barriers to mental health care were less common in parents or caregivers compared with adolescents. Embarrassment was not indicated as a barrier among any parent or caregiver; however, cost was the most common parent- or

guardian-reported barrier. There were no reductions in parent- or guardian-reported cognitive barriers to mental health care from baseline to the 4-week follow-up.

Mental Health Care

At 4 weeks, of the 42 adolescents who completed follow-up, 22 (52%) reported that they had attended an appointment with a mental health specialist. Of the 20 adolescents who had not attended a mental health appointment, 1 (5%) reported that it was scheduled for a future date, 2 (10%) reported that the primary care site did not have the ability to help them schedule a mental health care appointment, and 3 (15%) reported that they were no longer interested in mental health care.

Discussion

Principal Findings

In phase 1, we found evidence that adolescents with depression and their parents or caregivers face numerous barriers to receiving mental health care, among which cognitive barriers related to stigma and belief in the utility of mental health care play a significant role. We also found evidence that a program like T2C incorporating daily negative affect check-ins for adolescents and (for parent or caregivers) daily prompts to talk about mental health care with their child had the potential to reduce some barriers by being easy, relatable, and simpler than talking on the phone. In phase 2, we found evidence that T2C is used at high rates by adolescents, has high usability ratings, and may reduce perceived barriers over time.

Regarding adolescent engagement, we found that all completed onboarding, including mental health barrier classification, and there were high opt-in rates for continued daily negative affect check-ins beyond the first week. For adolescents, we believe that the brevity of daily interactions and the use of SMS text messaging contributed to these high engagement rates. Regarding parent or caregiver engagement, we found that although less than three-quarters of parents consented to enroll in T2C, all who did completed onboarding, including completion of mental health barrier classification. Not surprisingly, for adolescents with parents who refused enrollment, we found lower rates of mental health care initiation at follow-up. Future studies should seek to understand why certain parents or caregivers defer enrollment and ways to improve parental or caregiver enrollment in help-seeking interventions.

Usability ratings and qualitative feedback reinforced engagement findings. Overall, adolescents and parents or caregivers were satisfied with the T2C program, found the message content clear, and did not need to learn things before operation. In addition, around three-quarters of adolescents and parents or caregivers would recommend the T2C program to someone in need. From qualitative feedback, in addition to being seen as easy to use, the T2C program—especially the daily negative affect check-ins—seemed to provide some feeling of being cared for. The limitation of automated programs not being able to provide truly personalized support was brought up, suggesting that efficient programs that supplement *bot* with counselor communications should be considered [20].

At the 4-week follow-up, most adolescents reported either having initiated mental health care or pending scheduling,

leaving only 15% (3/20) of adolescents still not interested in mental health care. On the one hand, this seems to be a higher rate of mental health care initiation than the 30% reported in the literature. On the other hand, this indicates that additional program components or strategies are needed to close this gap. At baseline, almost three-quarters of parents or caregivers reported already making a mental health care appointment, and more than half of those remaining were ready to make an appointment. For these dyads, the T2C program functioned essentially as a scheduling assistant. For parents or caregivers who were not ready, the program sought to reduce cognitive barriers and promote dialog with their child about mental health care. Additional features may include strategies to overcome structural barriers to mental health help-seeking, including transportation and virtual visits.

Mechanistically, we found some evidence that the T2C program facilitates reductions in adolescent cognitive barriers to mental health care from baseline to 4 weeks. The greatest reduction (47% to 37%) was observed for concerns around embarrassment, with smaller reductions regarding privacy and disclosure. We speculate that the daily negative affect check-ins with feedback may have made adolescents more comfortable sharing mental health symptoms, which in turn helped reduce concerns about embarrassment. Future intervention design could consider enhancing this effect by priming the adolescent-provider relationship before an in-person visit with asynchronous communication.

Limitations and Strengths

This study has several limitations. First, we recruited mostly female White youth; therefore, the findings may not be valid in males or racial or ethnic minorities. Second, we did not measure objective mental health care initiation but used self-reports from adolescents. Third, we did not capture participants' specific mental health or substance use diagnoses in addition to depression, limiting the understanding of severity and breadth of mental health care needs. We note several strengths of our study. First, we were able to recruit youth and parent or caregiver dyads, as both are critical in understanding barriers to mental health care initiation. Second, we designed and evaluated the intervention through multistage user feedback, incorporating mixed methods. Third, in the phase-2 open trial, we examined multiple sources of outcome data within each participant to *triangulate* usability findings.

Conclusions

In conclusion, this two-phase study demonstrates the need to engage both adolescents and parents or caregivers in interventions aimed at overcoming barriers to mental health care initiation and the potential utility of digital strategies focused on cognitive and motivational barriers to help-seeking to do so. There is an urgent need for evidence-based help-seeking programs for youth with depression, and a program like T2C, if found to be effective in a larger trial, could fill a needed gap.

Acknowledgments

This work was supported by the National Institutes of Health grant 1P50MH115838 (principal investigator: DB).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample messaging.

[[DOCX File , 17 KB - formative_v5i11e30580_app1.docx](#)]

References

1. Ghandour RM, Sherman LJ, Vladutiu CJ, Ali MM, Lynch SE, Bitsko RH, et al. Prevalence and treatment of depression, anxiety, and conduct problems in US children. *J Pediatr* 2019 Mar;206:256-267 [[FREE Full text](#)] [doi: [10.1016/j.jpeds.2018.09.021](https://doi.org/10.1016/j.jpeds.2018.09.021)] [Medline: [30322701](#)]
2. Wang PS, Berglund PA, Olfson M, Kessler RC. Delays in initial treatment contact after first onset of a mental disorder. *Health Serv Res* 2004 Apr;39(2):393-315 [[FREE Full text](#)] [doi: [10.1111/j.1475-6773.2004.00234.x](https://doi.org/10.1111/j.1475-6773.2004.00234.x)] [Medline: [15032961](#)]
3. Chisolm DJ, Klima J, Gardner W, Kelleher KJ. Adolescent behavioral risk screening and use of health services. *Adm Policy Ment Health* 2009 Nov;36(6):374-380. [doi: [10.1007/s10488-009-0245-8](https://doi.org/10.1007/s10488-009-0245-8)] [Medline: [19777336](#)]
4. Asarnow JR, Jaycox LH, Duan N, LaBorde AP, Rea MM, Murray P, et al. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. *J Am Med Assoc* 2005 Jan 19;293(3):311-319. [doi: [10.1001/jama.293.3.311](https://doi.org/10.1001/jama.293.3.311)] [Medline: [15657324](#)]
5. Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. *J Am Acad Child Adolesc Psychiatry* 2019 Jan;58(1):72-79. [doi: [10.1016/j.jaac.2018.07.896](https://doi.org/10.1016/j.jaac.2018.07.896)] [Medline: [30577941](#)]
6. Bridge JA, Goldstein TR, Brent DA. Adolescent suicide and suicidal behavior. *J Child Psychol Psychiatry* 2006;47(3-4):372-394. [doi: [10.1111/j.1469-7610.2006.01615.x](https://doi.org/10.1111/j.1469-7610.2006.01615.x)] [Medline: [16492264](#)]
7. Barker G, Olukoya A, Aggleton P. Young people, social support and help-seeking. *Int J Adolesc Med Health* 2005;17(4):315-335. [doi: [10.1515/ijamh.2005.17.4.315](https://doi.org/10.1515/ijamh.2005.17.4.315)] [Medline: [16445071](#)]
8. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry* 2010 Dec 30;10:113 [[FREE Full text](#)] [doi: [10.1186/1471-244X-10-113](https://doi.org/10.1186/1471-244X-10-113)] [Medline: [21192795](#)]
9. Gulliver A, Griffiths KM, Christensen H, Brewer JL. A systematic review of help-seeking interventions for depression, anxiety and general psychological distress. *BMC Psychiatry* 2012 Jul 16;12:81 [[FREE Full text](#)] [doi: [10.1186/1471-244X-12-81](https://doi.org/10.1186/1471-244X-12-81)] [Medline: [22799879](#)]
10. Xu Z, Huang F, Kösters M, Staiger T, Becker T, Thornicroft G, et al. Effectiveness of interventions to promote help-seeking for mental health problems: systematic review and meta-analysis. *Psychol Med* 2018 Dec;48(16):2658-2667. [doi: [10.1017/S0033291718001265](https://doi.org/10.1017/S0033291718001265)] [Medline: [29852885](#)]
11. Patafio B, Miller P, Baldwin R, Taylor N, Hyder S. A systematic mapping review of interventions to improve adolescent mental health literacy, attitudes and behaviours. *Early Interv Psychiatry* 2021 Jan 03:13109. [doi: [10.1111/eip.13109](https://doi.org/10.1111/eip.13109)] [Medline: [33393231](#)]
12. Kauer SD, Mangan C, Sancil L. Do online mental health services improve help-seeking for young people? A systematic review. *J Med Internet Res* 2014 Mar 04;16(3):e66 [[FREE Full text](#)] [doi: [10.2196/jmir.3103](https://doi.org/10.2196/jmir.3103)] [Medline: [24594922](#)]
13. Peyton D, Hiscock H, Sciberras E. Do digital health interventions improve mental health literacy or help-seeking among parents of children aged 2-12 years? A scoping review. *Stud Health Technol Inform* 2019 Aug 08;266:156-161. [doi: [10.3233/SHTI190788](https://doi.org/10.3233/SHTI190788)] [Medline: [31397317](#)]
14. Rosenstock IM. Why people use health services. *Milbank Mem Fund Q* 1966 Jul;44(3):94-127. [doi: [10.2307/3348967](https://doi.org/10.2307/3348967)] [Medline: [5967464](#)]
15. Wang C, Bakhet M, Roberts D, Gnani S, El-Osta A. The efficacy of microlearning in improving self-care capability: a systematic review of the literature. *Public Health* 2020 Sep;186:286-296. [doi: [10.1016/j.puhe.2020.07.007](https://doi.org/10.1016/j.puhe.2020.07.007)] [Medline: [32882481](#)]
16. Schneiders J, Nicolson NA, Berkhof J, Feron FJ, van Os J, deVries MW. Mood reactivity to daily negative events in early adolescence: relationship to risk for psychopathology. *Dev Psychol* 2006 May;42(3):543-554. [doi: [10.1037/0012-1649.42.3.543](https://doi.org/10.1037/0012-1649.42.3.543)] [Medline: [16756444](#)]
17. Wisdom JP, Clarke GN, Green CA. What teens want: barriers to seeking care for depression. *Adm Policy Ment Health* 2006 Mar;33(2):133-145 [[FREE Full text](#)] [doi: [10.1007/s10488-006-0036-4](https://doi.org/10.1007/s10488-006-0036-4)] [Medline: [16489480](#)]
18. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum Comput Interact* 1995;7(1):57-78. [doi: [10.1080/10447319509526110](https://doi.org/10.1080/10447319509526110)]

19. Kuhl J, Jarkon-Horlick L, Morrissey RF. Measuring barriers to help-seeking behavior in adolescents. *J Youth Adolesc* 1997 Dec;637-650 [[FREE Full text](#)] [doi: [10.1023/a:1022367807715](https://doi.org/10.1023/a:1022367807715)]
20. 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/)]

Abbreviations

T2C: Text To Connect

Edited by J Torous; submitted 20.05.21; peer-reviewed by E Burner, J McCusker; accepted 12.08.21; published 09.11.21.

Please cite as:

Suffoletto B, Goldstein T, Brent D

A Text Message Intervention for Adolescents With Depression and Their Parents or Caregivers to Overcome Cognitive Barriers to Mental Health Treatment Initiation: Focus Groups and Pilot Trial

JMIR Form Res 2021;5(11):e30580

URL: <https://formative.jmir.org/2021/11/e30580>

doi: [10.2196/30580](https://doi.org/10.2196/30580)

PMID: [34751665](https://pubmed.ncbi.nlm.nih.gov/34751665/)

©Brian Suffoletto, Tina Goldstein, David Brent. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 09.11.2021. 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, Population-Based Cirrhosis Identification and Management System for Improving Cirrhosis Care: Qualitative Formative Evaluation

Sarah J Javier¹, PhD; Justina Wu¹, MPH; Donna L Smith², MEd; Fasiha Kanwal^{2,3,4}, MD; Lindsey A Martin⁵, PhD; Jack Clark⁶, PhD; Amanda M Midboe^{1,7}, PhD

¹Center for Innovation to Implementation (Ci2i), VA Palo Alto Health Care System, Menlo Park, CA, United States

²Center for Innovations in Quality, Effectiveness and Safety (IQeSt), Michael E. DeBakey VA Medical Center, Houston, TX, United States

³Section of Gastroenterology and Hepatology, Baylor College of Medicine, Houston, TX, United States

⁴Section of Health Services Research, Baylor College of Medicine, Houston, TX, United States

⁵National Institute of Environmental Health Sciences (NIEHS), Division of Extramural Research & Training, Population Health Branch, Research Triangle Park, NC, United States

⁶Department of Health Law, Policy, and Management, School of Public Health, Boston University, Boston, MA, United States

⁷Stanford School of Medicine, Stanford, CA, United States

Corresponding Author:

Sarah J Javier, PhD

Center for Innovation to Implementation (Ci2i)

VA Palo Alto Health Care System

795 Willow Road (152-MPD)

Menlo Park, CA, 94025

United States

Phone: 1 6504935000

Email: Sarah.Javier@va.gov

Abstract

Background: Cirrhosis, or scarring of the liver, is a debilitating condition that affects millions of US adults. Early identification, linkage to care, and retention of care are critical for preventing severe complications and death from cirrhosis.

Objective: The purpose of this study is to conduct a preimplementation formative evaluation to identify factors that could impact implementation of the Population-Based Cirrhosis Identification and Management System (P-CIMS) in clinics serving patients with cirrhosis. P-CIMS is a web-based informatics tool designed to facilitate patient outreach and cirrhosis care management.

Methods: Semistructured interviews were conducted between January and May 2016 with frontline providers in liver disease and primary care clinics at 3 Veterans Health Administration medical centers. A total of 10 providers were interviewed, including 8 physicians and midlevel providers from liver-related specialty clinics and 2 primary care providers who managed patients with cirrhosis. The Consolidated Framework for Implementation Research guided the development of the interview guides. Inductive consensus coding and content analysis were used to analyze transcribed interviews and abstracted coded passages, elucidated themes, and insights.

Results: The following themes and subthemes emerged from the analyses: outer setting: needs and resources for patients with cirrhosis; inner setting: readiness for implementation (subthemes: lack of resources, lack of leadership support), and implementation climate (subtheme: competing priorities); characteristics of individuals: role within clinic; knowledge and beliefs about P-CIMS (subtheme: perceived and realized benefits; useful features; suggestions for improvement); and perceptions of current practices in managing cirrhosis cases (subthemes: preimplementation process for identifying and linking patients to cirrhosis care; structural and social barriers to follow-up). Overall, P-CIMS was viewed as a powerful tool for improving linkage and retention, but its integration in the clinical workflow required leadership support, time, and staffing. Providers also cited the need for more intuitive interface elements to enhance usability.

Conclusions: P-CIMS shows promise as a powerful tool for identifying, linking, and retaining care in patients living with cirrhosis. The current evaluation identified several improvements and advantages of P-CIMS over current care processes and

provides lessons for others implementing similar population-based identification and management tools in populations with chronic disease.

(JMIR Form Res 2021;5(11):e27748) doi:[10.2196/27748](https://doi.org/10.2196/27748)

KEYWORDS

cirrhosis; informatics; care coordination; implementation; Consolidated Framework for Implementation Research (CFIR); quality improvement

Introduction

Background

Over 4.5 million US adults were diagnosed with liver disease in 2018, and an estimated 44,358 died of chronic liver disease in 2019 [1,2]. Recent estimates indicate that the total annual health care cost of advanced liver disease or cirrhosis is close to US \$9.5 billion [3]. Cirrhosis may be accompanied by a host of complications that require a complex, patient-centered treatment plan, including biannual surveillance visits [4,5]. However, patients with cirrhosis are often not diagnosed by their primary care provider or, alternatively, are not referred to specialty care after diagnosis [6]. Owing to the complexity of cirrhosis treatment, a systematic approach to cirrhosis management is critical to preventing further complications and the risk of death.

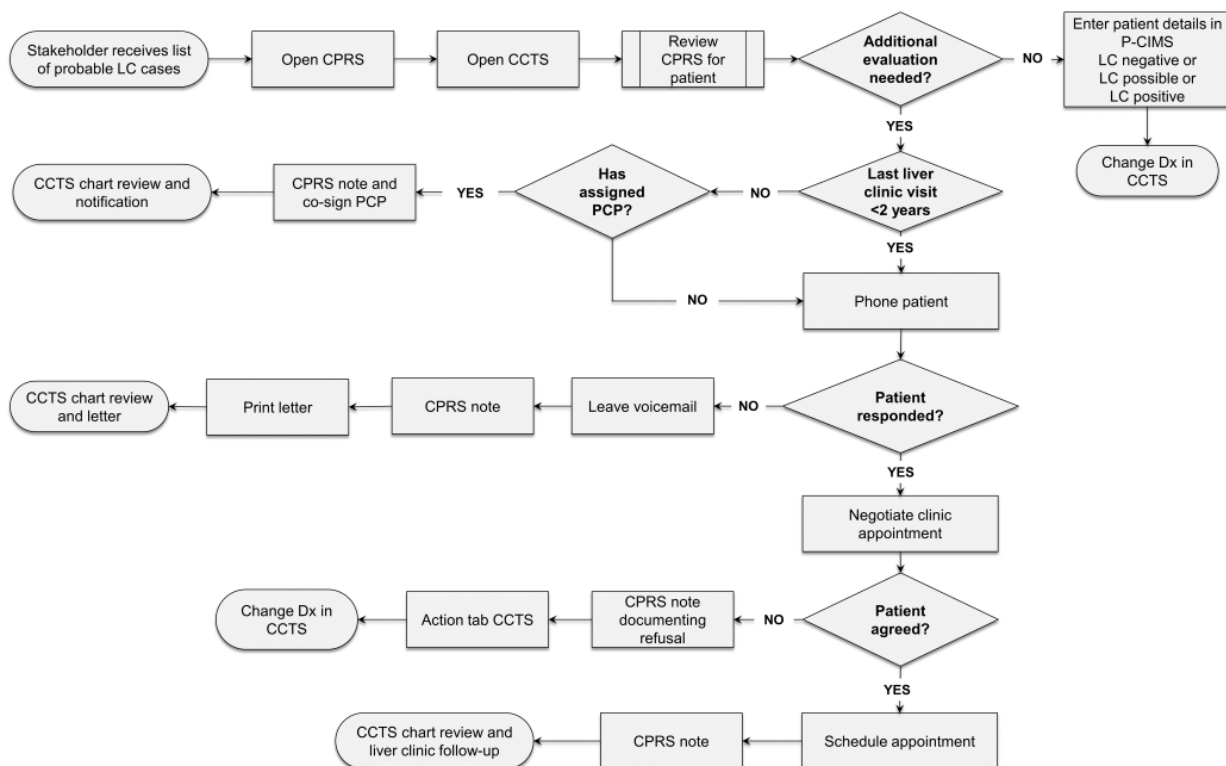
Health informatics tools have shown promise in the management of complex diseases and can enhance the continuity of care

[7-10]. However, it is often challenging to incorporate these tools into existing clinical workflows, with implementation success dependent on ongoing usability assessments [11,12]. In view of the high morbidity and mortality associated with cirrhosis, we conducted a preimplementation formative evaluation to identify facilitators and barriers to implementation of the Population-Based Cirrhosis Identification and Management System (P-CIMS) [13].

P-CIMS

P-CIMS, a secure, web-based informatics tool for providers, is designed to ensure efficient cirrhosis case identification in clinical populations, facilitate coordination of care for known cases, and prevent loss to follow-up [13]. Figure 1 presents the overall process workflow chart for P-CIMS. In Figure 1, we describe a scenario in which a provider receives a list of probable cirrhosis cases and must confirm cirrhosis diagnosis before tracking them in P-CIMS.

Figure 1. P-CIMS process workflow map (adapted from Kanwal et al [13]). CPRS: Computerized Patient Record System; CCTS: cancer case tracking system; Dx: diagnosis; LC: liver cirrhosis; P-CIMS: Population-Based Cirrhosis Identification and Management System; PCP: primary care provider.



Development and Testing

The precursor to P-CIMS, the Cancer Care Tracking System (CCTS), was developed by an interdisciplinary team of clinicians, programmers, and informatics experts in 2007 [9]. CCTS was a tool that consolidated local Veterans Health Administration (VHA) electronic medical record (EMR) data into a dashboard on which providers could view abnormal lung or liver cancer images and then track the clinical steps required to reach a definitive diagnosis and treatment plan.

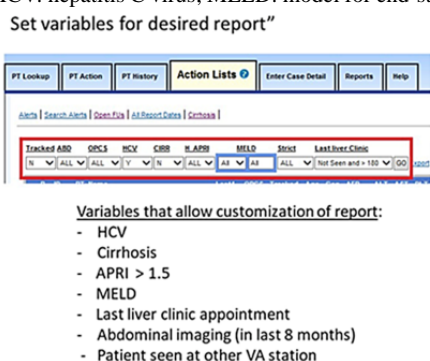
In 2015, P-CIMS was developed using CCTS as a base platform and then improved over 9 months [13]. P-CIMS data elements were validated using the full EMR from the VHA National Corporate Data Warehouse as the reference standard. The beta version described by Kanwal et al [13] was tested between 2015 and 2017 at a site with over 3800 patients diagnosed with cirrhosis, according to the EMR data. Results from the beta test indicated that the use of P-CIMS resulted in 30% of probable cases being referred to liver specialty care. Findings from the beta version testing version of P-CIMS have been published elsewhere [13].

Functional Components

There are 2 main functional components of P-CIMS:

1. Generation of lists of patients who are likely to have undiagnosed cirrhosis and patients with diagnosed cirrhosis who have been lost to follow-up using EMRs; and
2. Task management features designed to facilitate retention in care and help monitor whether patients complete recommended surveillance testing.

Figure 2. Static screenshots from the Population-Based Cirrhosis Identification and Management System, adapted from Kanwal et al [13] (top 2 images: generation of probable cases list; bottom 2 images: cirrhosis tracking). APRI: aminotransferase to platelet ratio index; FIB4: fibrosis-4; HCC: hepatocellular carcinoma; HCV: hepatitis C virus; MELD: model for end-stage liver disease; MRI: magnetic resonance imaging; VA: Department of Veterans Affairs.



Generation of Lists of Probable Cirrhosis Cases

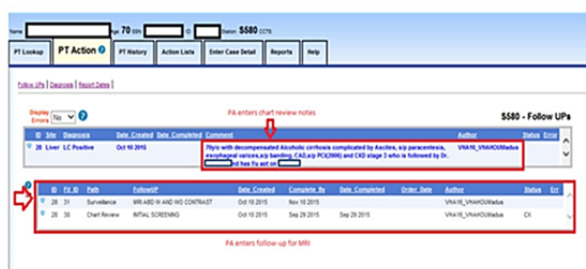
The P-CIMS report generator allows providers to generate lists of patients who are either likely to have undiagnosed cirrhosis or who have diagnosed cirrhosis but who have been lost to follow-up (henceforth, referred to as *probable cases*). Providers can customize these lists based on the following criteria:

1. Any outpatient or inpatient encounters in the last 3 years in which the patient had at least one cirrhosis diagnosis, as designated by the validated *International Classification of Diseases, Ninth Revision* and *Tenth Revision* codes [14];
2. Possible cirrhosis, defined as either aspartate aminotransferase to platelet ratio index >2.0, or Fibrosis-4 index >3.24 in patients with an active hepatitis C virus (HCV) infection.
3. Last visit to a liver or HCV clinic is >180 days earlier (ie, more than 6 months ago)

Probable cases must be confirmed by their provider via chart review for each patient on the list to confirm whether the patient needs to be coded in P-CIMS for cirrhosis surveillance tracking (Figure 1).

Cirrhosis Tracking and Management

The second function of P-CIMS is to track and manage all aspects of ongoing cirrhosis care for probable cases confirmed via chart review. Providers can track patient compliance with recommended consultations, screenings, surveillance testing, imaging, follow-up visits, or referral for transplant evaluation (Figure 2 shows static screenshots from P-CIMS).



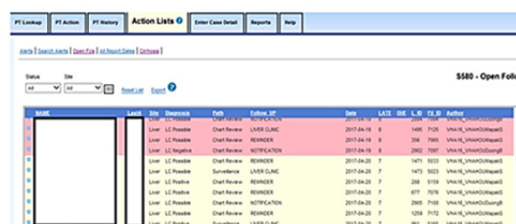
The clinician entered an MRI for HCC surveillance
Assigns a date (for when to expect an MRI)

Report criteria requested

- HCV+
- High FIB4 (≥3.25)
- Not seen in Liver/HCV clinic in the last 6 months (last visit >180 days prior or never seen)



Reports can be sorted by clicking the column header labels



Patient appears on list of all patients being followed.
List is color coded (green = due within 4 weeks; yellow = due now; red = overdue)
Follow-up tasks can be entered for ANY future date

Purpose of This Study

The purpose of this study is to conduct a preimplementation formative evaluation using qualitative interviews to identify barriers and facilitators of P-CIMS implementation across 3 testing sites, including at the beta testing site.

Methods

Setting

We conducted a qualitative formative evaluation with key stakeholders at 3 high-volume Department of Veterans Affairs (VA) medical centers in the Sierra Pacific and Southern Central United States (N=10), including a high-volume site at which beta testing occurred. The other 2 sites were chosen based on the criteria for successful P-CIMS implementation. These criteria included the following: (1) established relationships between external facilitators (FK, DLS, and AMM) and key stakeholders; (2) internal facilitators, such as primary care and liver specialty clinical leadership support; and (3) liver clinic patient volume. The sites that were chosen were known to have a higher than average liver clinic volume.

Sample

Although the key purpose of this study was to identify potential implementation barriers and facilitators through qualitative interviews, we were also interested in examining usability. As such, we centered on interviews with specific end-users who would be using the tool for clinical care; we followed sample size recommendations from the empirical literature. Guidance by Turner, Lewis, and Nielsen maintains that a minimum of 3 to 5 participants are sufficient for a usability test with the caveat that running additional subjects during the same test is unlikely to reveal new information [15]. Further, other guidance from the quality improvement literature maintains that code saturation can be reached with as few as 9 interviews [16]. Our final sample comprised 10 providers, including 8 physicians and midlevel providers from liver-related specialty clinics and 2 primary care providers who managed patients with cirrhosis.

Design

Interview Guide Development and Testing

We created a semistructured interview guide based on the Consolidated Framework for Implementation Research (CFIR), a metatheoretical framework that describes a set of constructs organized within 5 domains [17]. These constructs allow implementation scientists to systematically assess barriers and facilitators to implementation, including contextual factors relevant to multisite implementation projects. The following CFIR domains were chosen for P-CIMS evaluation: outer, inner, and individual characteristics.

The interview guide was developed, iteratively modified, and pilot-tested by 2 qualitative researchers with training in medical anthropology and public health (LAM and JC). The pilot interview participants were gastroenterologists from the P-CIMS beta test site and a gastroenterologist from a nontest site who had previously seen a demo version of P-CIMS. Feedback from these interviewees was used to refine the interview guide.

Semistructured Interviews

The qualitative team conducted semistructured interviews between January and May of 2016. Providers were asked about current practices for identifying and linking patients with cirrhosis to liver clinic specialty care. Interview topics included the perceived value and benefit of P-CIMS, usability, barriers to implementation, and opportunities to improve the tool. VHA privacy rules prevented live demonstration of the 2 nontest sites. For those interviews, a Microsoft PowerPoint demonstration, including deidentified static screenshots of P-CIMS, was shown to the interviewees. An expert was available during the interviews to answer questions about the tool.

At the beta test site, the interviews included content-specific usability questions. Four providers who had already used the beta version of P-CIMS were asked questions that explored how they had integrated P-CIMS into their clinical workflow, challenges in using the system, and suggestions for improvement (LAM and JW). Participants were asked to open P-CIMS and walk through how they used the system while noting down any features that were helpful or not helpful.

Analytic Approach

Content analysis is the primary approach used in the current evaluation [18]. Two analysts (LAM and JW) independently reviewed the transcripts and constructed the codes to describe the data. An iterative consensus process was used to draft a list of subcodes, exemplar quotes, and broader code categories. This list was then used to create a codebook consisting of a priori codes that emerged from the data (eg, *suggestions for improving P-CIMS*) and code definitions [18]. Individual codes were also deductively linked to CFIR constructs that guided the data collection. For example, *suggestions for improving P-CIMS* was linked to the CFIR construct *Inner Setting: Readiness for Implementation*. These constructs served as sensitizing concepts to guide coding, but still allowed for the identification of other salient themes in the data [18].

The codebook was applied for consensus-based coding of the remaining transcripts. Two coders (LAM and JW) met regularly to identify additional codes and refine existing code definitions, with any discrepancies resolved through negotiated consensus. One member of the qualitative team (JW) compiled coded passages into separate transcripts, which were used to derive subthemes. Subthemes were then compiled into matrices for each code, which included exemplar quotes and examples. For example, for the major code *suggestions for improving P-CIMS*, one subtheme was *adding patients to P-CIMS*, with descriptive examples, *automatic instead of manual entry*, and *need clear definition of cirrhosis*.

Several steps were taken to bolster the validity of data collection and analysis. First, 2 pilot interviews with the target users of P-CIMS were used to refine the initial codebook (See *Interview Guide Development and Testing* section). This is a step that is often skipped in qualitative research, but it allows investigators to address instrumentation and bias issues before actual data collection occurs [19]. Second, the qualitative team purposively sampled interview participants and interviewed only individuals who would provide the most appropriate and meaningful insights

on P-CIMS [20]. To that end, 20% (2/10) primary care providers were interviewed, and specialists (eg, gastroenterologists) comprised most of the sample (8/10, 80% providers). These specialists are the target end-users for P-CIMS, as they work most closely with patients who would benefit from the tool. In contrast, we found that interviews with the 2 primary care providers did not offer unique information in comparison to specialty care providers. Therefore, we chose not to interview primary care providers. Third, the qualitative team employed 2 types of triangulation of analytic findings: theoretical triangulation and investigator triangulation [21]. Theoretical triangulation, defined as the use of substantive theoretical lenses (ie, CFIR) to drive data collection and review research findings, provided a strong basis for the development of both the interview guide and the codebook. Investigator triangulation, or data analysis by more than 2 independent coders with experience in qualitative research (ie, LAM, JW, and DLS) ensured that concurrence for the final list of subthemes and descriptive examples was achieved. Data analyses were iterative and continued until thematic saturation was reached.

Ethics

The Stanford University Human Research Protection Program reviewed and approved this project as a quality improvement project. All participants provided verbal consent before being interviewed. For this type of study (ie, quality improvement), formal consent is not required.

Human Rights

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The Stanford University Human Research Protection Program reviewed and approved this project as a quality improvement project. The VA Central Institutional Review Board also approved project #14-11 Population-based identification and management of veterans with HCV cirrhosis. For this type of study (ie, quality improvement), formal consent was not required.

Informed Consent

All participants gave verbal consent before being interviewed.

Results

Sample

Nine interviews took place in person or over the phone and were audio-recorded and transcribed. One interviewee requested that the interview not be recorded, and extensive notes were taken for this interview.

CFIR Mapping

Table 1 presents the main findings mapped to the CFIR constructs of the outer and inner settings and individual characteristics of the individuals. Figure 3 presents the main findings mapped to these CFIR constructs in a descriptive figure.

Table 1. Consolidated framework for implementation research constructs in Population-Based Cirrhosis Identification and Management System implementation.

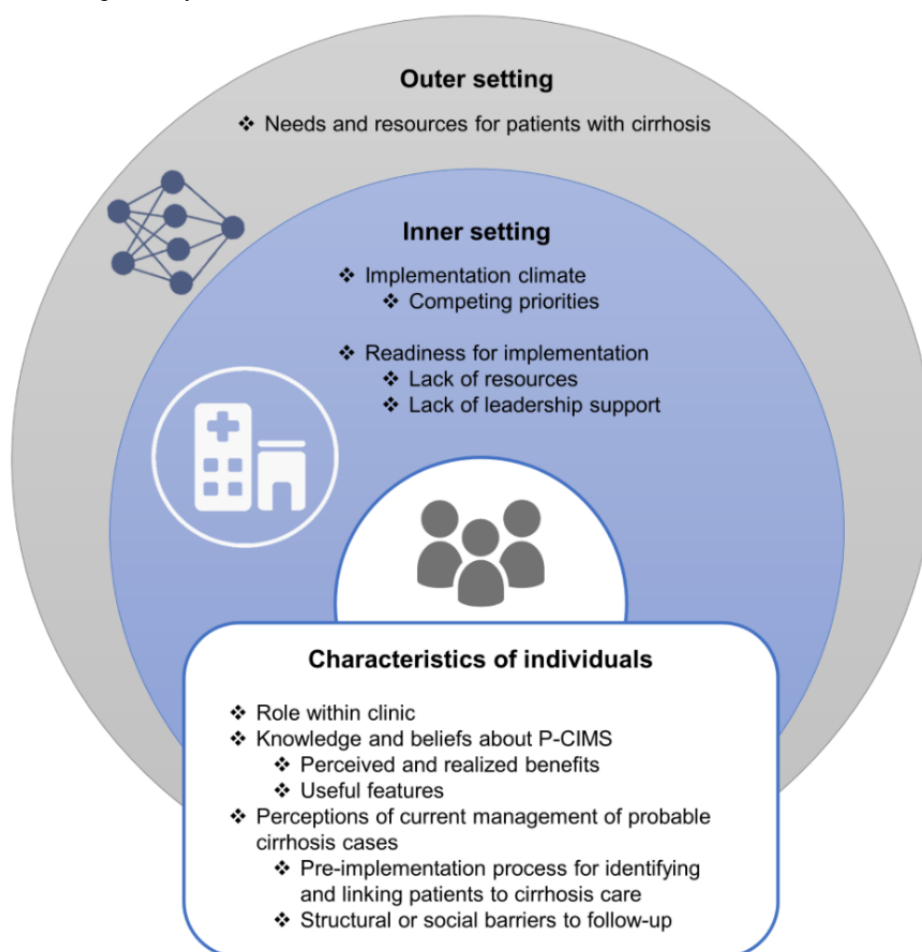
CFIR ^a construct	Example from evaluation	Illustrative quote
Outer setting domain—needs and resources for patients	Health informatics tools like P-CIMS ^b can enhance care continuity for patients with cirrhosis	“If patients are lost to follow-up for 2 years, they drop off our panel...we don’t currently have any way of...reaching out to them like the cirrhosis tracker would do.”
Inner setting domain—readiness for implementation	Providers said that lack of resources might prevent them from integrating P-CIMS in their clinical workflow	“We don’t necessarily have the time in our day-to-day duties to really do justice to what this tool really can do.”
Inner setting domain—implementation climate	Leadership is not as engaged with cirrhosis initiatives as other initiatives at their location	“...we have problems getting enough support for HCV ^c care...cirrhosis is not on the radar.”
Characteristics of individuals—role within clinic	Providers perceive that cirrhosis tracker implementation is not a part of their role	“If we had a very clear algorithm that the liver clinic could sign off on, then potentially the nurses could work through that but you’d have to work with the nursing service to their agreement that it was within their scope.”
Characteristics of individuals—knowledge and beliefs about the intervention	P-CIMS can be used to streamline continuity of care for patients with cirrhosis	“...it could very well change the landscape of hepatology, cirrhosis care as we know it. Anything that we can do that’s going to be innovative and...improve access and quality—typically is adopted.”
Characteristics of individuals—perceptions of current practices	Manual tracking systems for patients with cirrhosis are ineffective	“We used to do liver lesion monitoring where we manually entered patients into Excel on a monthly basis. This didn’t work well...we no longer do this, use the dashboard instead.”

^aCFIR: Consolidated Framework for Implementation Research.

^bP-CIMS: Population-Based Cirrhosis Identification and Management System.

^cHCV: hepatitis C virus.

Figure 3. Results of qualitative analysis mapped to Consolidated Framework for Implementation Research constructs. P-CIMS: Population-Based Cirrhosis Identification and Management System.



Characteristics of Individuals: Role Within Clinic

Overall, 50% (5/10) of the interviewees discussed how conflicting role expectations may hinder P-CIMS implementation. For example, liver clinic specialists expressed concern that primary care providers may not be as well versed in cirrhosis management as liver or infectious disease specialists. One provider said:

[The tool] makes sure that the referring providers are aware that these patients have risk factors for advanced fibrosis cirrhosis and underlying Hepatitis C and need to be seen in our service. [ID6]

Furthermore, some providers said that without an automated reminder system, scheduling staff may forget to contact patients for follow-up appointments. One provider said:

The liver clinic relies on automated CPRS callback and reschedule procedures. Patients who are due to return to clinic receive an automated phone reminder and an auto-generated appointment reminder card prior to the expected clinic visit date. Similar procedures are in place for when a patient is a "no-show." No one is assigned to track follow-up beyond the automated process. [ID8]

Characteristics of Individuals: Perceptions of Current Practices

Preimplementation Process for Identifying and Linking Patients to Cirrhosis Care

Overall, 60% (6/10) of the providers discussed the current processes for identifying and linking patients to cirrhosis care. Before P-CIMS implementation, liver clinic staff relied on referrals from primary care or infectious disease clinics for the identification of patients with cirrhosis. After a patient with cirrhosis was linked to the liver specialty clinic, some liver clinic providers created manual tracking systems to manage patient follow-up office visits and surveillance testing. However, as one provider explained, these tools could be cumbersome for the staff to maintain:

So the group of patients we used to track in the past are people who had abdominal imaging results...we actually have a list that we follow to make sure that they get the proper form of imaging...enter them, manually enter them into spreadsheets and then we actually make sure that every 6 months they have the imaging....But we had to stop doing that because we just don't have any manpower to be able to do that. [ID4]

Structural and Social Barriers to Follow-Up

Overall, 70% (7/10) of the providers cited various reasons for patients with cirrhosis not attending follow-up appointments or not being linked to care after their initial appointment. Some barriers are structural or social in nature. For example, even though the VA routinely sent automated telephone appointment reminders to patients, many did not keep their appointments. Providers cited lack of response as attributable to potential barriers such as homelessness, drug addiction, or lack of access to mail or a phone. Second, if a patient canceled an appointment, there was no alert system in the EMR to notify the provider or to remind the administrative staff that the patient needed rescheduling. A third reason for loss to follow-up, as reported by providers, was the characteristics of the patients themselves. For instance, patients in the early stages of cirrhosis may be asymptomatic. These patients may not monitor their condition closely and, consequently, may stop attending ultrasound and other appointments. For instance, one provider stated:

Everything can be perfect at our end, but he or she would not show up for [an appointment]. So education is the first and foremost thing for them to understand the gravity of the condition and the importance of follow-up even though they feel well. [ID1]

In addition, many patients with multiple comorbidities may not prioritize cirrhosis when other health issues cause immediate discomfort. Finally, providers reported that some patients may choose not to pursue any treatment for their liver conditions.

Inner Setting Barrier: Readiness for Implementation

Overall, 50% (5/10) of the interviews discussed clinical readiness for implementing P-CIMS. Overall, providers felt that P-CIMS would be best suited for use in a liver clinic or other specialties that regularly identified and managed patients with cirrhosis. Other providers cited barriers such as the lack of available time and number of staff needed to learn and use the tool, management of cirrhosis as not being part of primary care focus, and the worry that P-CIMS would add to the workload burden. For instance, one provider said:

I'm just a little bit nervous about any software that's going to require a lot of extra action. [Primary care] is not going to review 1200 patients to see who is a candidate for this. Providers are pretty burnt out. [ID5]

Participants acknowledged that it would take some time and training before some services fully embraced P-CIMS. Providers whose local decision-makers valued innovative tools and their potential to improve patient care felt that P-CIMS was likely to be adopted in their setting. However, the system would need complete buy-in from all providers as well as thoughtful planning about how to best integrate it into the current workflow.

Lack of Resources

Overall, 50% (5/10) of the providers pointed out the lack of adequate staff time and resources as the biggest obstacle to implementing and using P-CIMS. As one nurse practitioner noted:

We don't necessarily have the time in our day-to-day duties to really do justice to what this tool really can do. So finding the time, carving out the time, the uninterrupted time to be thorough in doing this is the issue. [ID6]

Participants said that P-CIMS needed a designated coordinator to coordinate and send reminders to primary care for follow-up. However, most interviewees stated that such a coordinator would need additional support. One provider said:

Either we have one person who is 100% dedicated to putting in all of the information regarding all of the patients that are being seen in the tracker and then put their follow-up times, date, and whatever things we need...Even that person may or may not be able to do all of the things that are necessary because the workload is going to be significant. [ID10]

In addition, the participants said that adding P-CIMS duties would slow down their productivity. Reports of chart review times, along with accompanying patient telephone contact activities, were closer to 12 to 13 minutes instead of 7 minutes, as originally estimated by the initial beta testers. Some providers also said that any new patients identified through P-CIMS would increase the demand for other hospital services such as radiology, procedures, and laboratory tests. To address this limitation, one option proposed by a provider was to distribute the P-CIMS duties among several staff members. They stated:

On our [Patient Aligned Care Team] team, we have an LVN and an RN and a pharmacist...I think all 3 of them, you know, have some time when I give them specific tasks, like dashboard tasks to do, I think they can do it. [ID12]

One caveat to this strategy is that it involves a high level of coordination to ensure that inefficiencies and redundancies do not occur.

Ideas varied about the best way to integrate the use of P-CIMS into the clinic workflow. Most interviewees agreed that the user would need uninterrupted time to use the P-CIMS accurately. With the current state of resources, providers said that uninterrupted time would most likely be after hours or on weekends. As an alternative approach, one provider suggested entering patient data into the tool at the end of a visit, when the patient's information is fresh in the provider's mind. They stated:

I think it probably would require anywhere from 3 to 5 minutes to complete [patient tracker] because you are fresh and then you are actually looking at the patient record and you've just completed the notes. [ID10]

Estimates varied across providers as to how often the tool would need to be beneficial.

Some interviewees said that the tool should be used once every 3 to 6 months, whereas others said that once per month use would be sufficient.

Lack of Leadership Support

A total of 40% (4/10) of the providers discussed the lack of leadership support for tools such as P-CIMS. One provider noted that cirrhosis was not a leadership priority at their location, and that buy-in for P-CIMS would require more leadership engagement. They said:

If it's a financially productive activity... [administration] might be willing to provide support. You're not gonna save money by getting more patients into hepatitis C treatment, so you know, it's sort of a bottom line for the administration....The motivation would have to be better patient care. It couldn't be a financial motivation. [ID12]

To achieve this, the provider recommended highlighting the tool's potential to improve patient care to site leadership.

Inner Setting Barrier: Implementation Climate

Competing Priorities

A total of 70% (7/10) providers discussed competing resources at their sites. The competing priorities were similar across the primary care providers at different sites. Providers stated that the treatment of chronic conditions, such as heart failure or diabetes, took precedence over cirrhosis. In addition, providers described having too many dashboard tools to use. For instance, one provider said:

We already have an opioid dashboard. We have an endocrine dashboard. We have diabetes, we have hypertension, and to be honest, almost nobody has time to actually look at them. [ID5]

Most of these tools are not regularly used by providers because of time constraints, or because they find it confusing and onerous to use multiple tools. Finally, as mentioned in most interviews, current staffing levels and clinic duties made it difficult to find time to incorporate a new tool into the current workflow.

Characteristics of Individuals: Knowledge and Beliefs

Perceived and Realized Benefits

Most participants (ie, 8/10, 80%) generally understood and valued P-CIMS's ability to link patients who are most likely to fall through the cracks or who are at the highest risk of cirrhosis to care. One provider said:

We [infectious disease clinic] are more focused on treating patients before they develop cirrhosis, so for me the focus will be Hepatitis C and Hepatitis B infected patients, identifying them before they have cirrhosis, and the tracker allows this to happen. [ID9]

Other providers felt that the P-CIMS's ability to identify these patients and help providers monitor and follow up with patients using reminders and task lists was its most valuable benefit. Some providers even saw the potential for P-CIMS to be used for other chronic conditions such as diabetes, high blood pressure, and heart failure. For instance, one provider said:

I think also it can be used to help in other disease processes, not just liver. Because it has the upside then of monitoring diabetes, monitoring your high blood pressure, possibly monitoring patients with COPD, CHF. [ID6]

Many found it more powerful and user-friendly than existing tools such as the clinical case registry in its ability to store, track, and organize information about patients.

Useful Features

Most providers (ie, 7/10, 70%) stated that certain P-CIMS features were useful. Providers found the ability to set reminders and track patient follow-up actions to be one of the most useful aspects of P-CIMS. Functions such as sorting capabilities and actions color-coded by priority were also cited as helpful. In addition, providers also liked the following features: radiology reports, calculated measures of cirrhosis risk (eg, model for end-stage liver disease score, Fibrosis-4 index for liver fibrosis, aspartate aminotransferase to platelet ratio index), primary care physician reminders to place consults, cancer tracking options, patient look-up, action lists, and capturing barriers to care. Suggestions for P-CIMS improvement

Suggestions for Improvement

All interviewees (ie, 10/10, 100%) provided suggestions for improving P-CIMS. Providers' suggestions for P-CIMS improvement were categorized into 4 themes: changes that would buffer against a slowdown of clinic productivity, changes to existing content, usability improvements, and potential new features. [Table 2](#) presents these suggestions based on theme.

Table 2. Provider suggestions for Population-Based Cirrhosis Identification and Management System improvement by theme.

Avoiding productivity slowdown	Changes to existing content	Usability improvements	Potential new features
<ul style="list-style-type: none"> • Use P-CIMS^a outside of clinic hours or on weekends • Use P-CIMS after completing patient notes 	<ul style="list-style-type: none"> • Provide relevant information in prominent areas of P-CIMS • Provide a clear definition of cirrhosis on P-CIMS to guide decision-making • Provide a guide for midlevel providers diagnosis cirrhosis 	<ul style="list-style-type: none"> • Include patient information (eg, telephone number) in the tool to bypass referring to the EHR^b • Autopopulate P-CIMS with EHR data • Implement an autosave feature to avoid loss of work • Incorporate tracking of clinical surveillance guideline requirements • Track use metrics (eg, the amount of time each user spends on P-CIMS) • Incorporate a filter feature where patients can be filtered by providers • Autopopulate follow-up dates 	<ul style="list-style-type: none"> • Add search feature for abbreviations or acronyms • Add at-a-glance tab • Add a pending tab • Add quick-access buttons • Add an indicator for a closed consultation

^aP-CIMS: Population-Based Cirrhosis Identification and Management System.

^bEHR: electronic health record.

Discussion

Principal Findings

A preimplementation formative evaluation using the CFIR domains of outer setting, inner setting, and characteristics of individuals identified several key benefits and barriers to implementing a health informatics tool for the management of patients with cirrhosis—P-CIMS—at 3 sites. While participants expressed overall interest in and appreciation for the potential value of P-CIMS, there was concern about the feasibility of implementation. Mainly, barriers in the inner setting—understaffing and workload—posed the biggest challenges in the implementation of P-CIMS.

During the formative evaluation process, providers discussed current strategies for linking patients with cirrhosis to liver clinic specialty care (characteristics of individuals: perceptions of current management of patients with cirrhosis). In the absence of P-CIMS, providers used manual spreadsheets and manual tracking systems to manage scheduled office visits, surveillance testing, and care. For some clinics, the integration of digital tools in addition to regular electronic health record use may be onerous and result in less efficient forms of documentation. For instance, the use of electronic health records to conduct clerical and administrative tasks has been shown to comprise nearly half of some clinicians' overall workdays [22]. Providers burdened with these tasks may seek out alternatives to P-CIMS to track patients with cirrhosis. Overreliance on nonelectronic health record documentation methods, such as noncoded text notes, has been shown in the past studies of health data [23]. In this evaluation, there were major concerns that adequate implementation of P-CIMS would ultimately slow down clinical productivity (inner setting: readiness for implementation). Participants suggested that having a coordinator may circumvent time barriers but perceived that their departments did not have the resources to hire someone to fulfill this role (outer setting: needs and resources for patients with cirrhosis). However,

participants (clinical staff) felt that the tool was useful and that they could use it outside their clinic.

The lack of resources to hire a coordinator was because of the lack of buy-in from leadership in their settings for use of clinical tools such as P-CIMS (Inner Setting: Readiness for Implementation). Providers at 2 locations reported that cirrhosis management was not currently a priority for their leadership, and the importance of cirrhosis was secondary to other competing priorities. Health systems have historically established chronic disease priorities based on the overall prevalence of chronic conditions across patients [24]. The apparent lack of leadership support indicates that health informatics tools such as P-CIMS may be more successfully implemented if they are prioritized and championed in the inner clinic setting, or if their importance for patient outcomes is emphasized to leadership.

A second version of the web-based tracking system, based on feedback from this formative evaluation, is under development. In part, this revised version is under development because preliminary results show that the first version of P-CIMS successfully facilitated linkage to liver clinic specialty care for approximately 30% of patients identified as possibly needing care and who were not already being seen in the clinic [13]. The second version will include key upgrades to P-CIMS, including features that facilitate identification and linkage to care of patients living with cirrhosis, as well as improvements in usability. To facilitate these improvements, data will first be pulled from a national EMR instead of a local VHA EMR. This will facilitate the spread of P-CIMS across VHAs nationally. Second, patient identification can be improved to facilitate linkage and retention in care. The new version will identify patients within 30 days of an inpatient diagnosis of cirrhosis, thus increasing the likelihood that liver clinic providers will follow up and retain these patients in care. This version will also identify patients with no formal diagnosis of cirrhosis in the national EMR data, but those with *International*

Classification of Diseases, Tenth Revision codes for complications of cirrhosis. These more strategic identification methods are another step in the process of linking patients, who might not otherwise be referred, with specialty care. The final change to patient identification will be to include patients with current or past evidence of positive HCV RNA at any point in VHA care (instead of a recent positive HCV RNA) with fibrosis-4 score >3.25 . This will enable providers to monitor hepatocellular carcinoma surveillance of patients who have attained virological cure from direct-acting antiviral treatments. Finally, the user interface will be streamlined to decrease the cognitive load of the end user, streamlining multiple tabs into fewer tabs while providing the necessary information to make clinical decisions. Changes to the second version address several limitations of the initial version. Developers and implementers of similar systems should consider these changes, particularly using a streamlined interface for a busy clinic, for future work in the management of chronic diseases such as cirrhosis.

Limitations

Although we believe this study is a valuable contribution to the literature examining the implementation of web-based health informatics tools, we do want to address some limitations. First, the study sample size was smaller than that recommended by some standards for formative research evaluations [20]; however, we followed best practices outlined for human factors when evaluating informatics tools [15,25,26]. We also reached thematic saturation related to the identification of the key factors impacting P-CIMS implementation [27]. Within these parameters, our sample size was appropriate, given that the purpose of this study was to identify factors that impacted P-CIMS implementation among a purposive sample of providers of cirrhosis and liver care.

A second limitation was that the provider and care team interviewees in the current sample may have been biased in their perception of P-CIMS as a tool. In the preimplementation phase, sites were chosen based on their perceived ability to implement a cirrhosis management tool. Key characteristics of successful site candidates included established relationships with potential liver clinic champions in the inner setting and external factors that facilitated the implementation of such tools. Furthermore, only 20% (2/10) of the primary care providers were included in the sample. Their perspectives may not have been representative of all VHA primary care providers serving patients with cirrhosis. Barriers and facilitators of implementation may differ in sites without these facilitators. Therefore, these findings may have limited generalizability to both other VA settings and non-VA settings, since data were

collected from only 3-VA facilities. These results may also not be generalizable to health care settings that do not have the benefit of a national integrated system of data from EMRs.

Finally, we understand that there is a small likelihood that adverse events may occur with the use of P-CIMS, but our findings from this and the larger evaluation do not indicate that the risk of using the tool is higher than the risks associated with current clinical practices used to track and manage patients with cirrhosis. As reflected in the extant literature, many probable cirrhosis cases are not diagnosed with cirrhosis or seen in specialty liver care. P-CIMS streamlined patient identification and created pathways to improve the coordination and continuity of care [13].

Impact

P-CIMS is the first informatics tool to leverage EMR data to improve the quality of care for patients with cirrhosis. This tool can be used to inform the design of other clinical informatics tools for a variety of chronic disease conditions that require close tracking and management. However, as is apparent from our qualitative analyses, tools need to be adapted to meet the needs of understaffed clinics or clinics with high workloads. Furthermore, these findings should be used to translate the implementation of health informatics tools for tracking and monitoring chronic diseases in different settings, using tailored implementation strategies as necessary. For instance, future implementation studies could use implementation frameworks such as CFIR, the capability, opportunity, and motivation model, and the behavior change wheel to both: (1) assess major barriers within specific settings, and (2) map these barriers to implementation strategies [28].

Conclusions

In future studies, patients living with chronic diseases such as cirrhosis should be engaged as key stakeholders during the implementation phase of a novel health informatics tool. It is important to elicit feedback on how health informatics tools can either facilitate or hinder chronic disease management. In this study, lack of resources, lack of support, and competing interests were found to be major barriers to the implementation of this informatics tool. Although these barriers may also be found in other settings, it is important to use implementation frameworks to conduct formative evaluations and adapt strategies that would best translate in those settings. The key element of a formative evaluation in this context is that it enables users' perceptions of the innovation, in the context of their clinical practices, to be incorporated into the design of a successful, sustainable solution to practice problems.

Acknowledgments

This study was supported by the Department of Veterans Affairs (VA) QUERI Program (QUE 15-284), VA HIV, Hepatitis C, and Related Conditions Program, and VA National Center for Patient Safety. This work was also supported in part by the Veterans Health Administration Center for Innovations in Quality, Effectiveness and Safety (CIN 13-413): Michael E DeBakey VA Medical Center, Houston Texas, and the Center for Gastrointestinal Development, Infection, and Injury (NIDDK P30 DK 56338). The views expressed in this study are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or of the United States government.

Authors' Contributions

SJJ wrote the manuscript with support from AMM, DLS, JW, JC, and LAM. LAM and JW conducted the qualitative analyses. FK, AMM, and DLS established the partnerships. LAM and JW conducted the interviews.

Conflicts of Interest

None declared.

References

1. Chronic liver disease and cirrhosis. FastStats CDC. URL: <https://www.cdc.gov/nchs/fastats/liver-disease.htm> [accessed 2021-10-01]
2. About underlying cause of death, 1999-2019. Centers for Disease Control and Prevention. URL: <http://wonder.cdc.gov/ucd-icd10.html> [accessed 2021-10-01]
3. Brewster C. Alcoholic cirrhosis imposes financial burden on US healthcare. HCP Live. 2018. URL: <https://www.hcplive.com/view/alcoholic-cirrhosis-imposes-financial-burden-on-us-healthcare> [accessed 2021-10-01]
4. Kanwal F, Tapper EB, Ho C, Asrani SK, Ovchinsky N, Poterucha J, et al. Development of quality measures in cirrhosis by the Practice Metrics Committee of the American Association for the study of liver diseases. *Hepatology* 2019 Apr;69(4):1787-1797. [doi: [10.1002/hep.30489](https://doi.org/10.1002/hep.30489)] [Medline: [30586188](https://pubmed.ncbi.nlm.nih.gov/30586188/)]
5. Mellinger JL, Volk ML. Multidisciplinary management of patients with cirrhosis: a need for care coordination. *Clin Gastroenterol Hepatol* 2013 Mar;11(3):217-223 [FREE Full text] [doi: [10.1016/j.cgh.2012.10.040](https://doi.org/10.1016/j.cgh.2012.10.040)] [Medline: [23142204](https://pubmed.ncbi.nlm.nih.gov/23142204/)]
6. Kanwal F, Volk M, Singal A, Angeli P, Talwalkar J. Improving quality of health care for patients with cirrhosis. *Gastroenterology* 2014 Dec;147(6):1204-1207. [doi: [10.1053/j.gastro.2014.10.029](https://doi.org/10.1053/j.gastro.2014.10.029)] [Medline: [25449026](https://pubmed.ncbi.nlm.nih.gov/25449026/)]
7. Armstrong W, Del Rio C. Patient tracking as a tool to improve retention in care: is the juice worth the squeeze? *Clin Infect Dis* 2017 Jun 01;64(11):1555-1556. [doi: [10.1093/cid/cix196](https://doi.org/10.1093/cid/cix196)] [Medline: [28329039](https://pubmed.ncbi.nlm.nih.gov/28329039/)]
8. Haroon S, Adab P, Riley RD, Fitzmaurice D, Jordan RE. Predicting risk of undiagnosed COPD: development and validation of the TargetCOPD score. *Eur Respir J* 2017 Jun;49(6):1602191 [FREE Full text] [doi: [10.1183/13993003.02191-2016](https://doi.org/10.1183/13993003.02191-2016)] [Medline: [28642308](https://pubmed.ncbi.nlm.nih.gov/28642308/)]
9. Taddei TH, Hunnibell L, DeLorenzo A, Rosa M, Connery D, Vogel D, et al. EMR-linked cancer tracker facilitates lung and liver cancer care. *J Clin Oncol* 2012 Dec 01;30(34):77. [doi: [10.1200/jco.2012.30.34_suppl.77](https://doi.org/10.1200/jco.2012.30.34_suppl.77)]
10. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. *J Med Internet Res* 2018 Feb 05;20(2):e41 [FREE Full text] [doi: [10.2196/jmir.8793](https://doi.org/10.2196/jmir.8793)] [Medline: [29402759](https://pubmed.ncbi.nlm.nih.gov/29402759/)]
11. Kvedar J, Fogel A. Why real-world results are so challenging for digital health. NEJM Group. 2017. URL: <https://catalyst.nejm.org/doi/full/10.1056/CAT.17.0453> [accessed 2021-10-01]
12. Workflow tool examples. Agency for Healthcare Research and Quality. URL: <https://healthit.ahrq.gov/health-it-tools-and-resources/evaluation-resources/workflow-assessment-health-it-toolkit/examples> [accessed 2021-10-01]
13. Kanwal F, Mapaskhi S, Smith D, Taddei T, Hussain K, Madu S, et al. Implementation of a population-based cirrhosis identification and management system. *Clin Gastroenterol Hepatol* 2018 Aug;16(8):1182-6.e2 [FREE Full text] [doi: [10.1016/j.cgh.2018.01.041](https://doi.org/10.1016/j.cgh.2018.01.041)] [Medline: [29803805](https://pubmed.ncbi.nlm.nih.gov/29803805/)]
14. Kramer J, Davila J, Miller E, Richardson P, Giordano T, El-Serag H. The validity of viral hepatitis and chronic liver disease diagnoses in Veterans Affairs administrative databases. *Aliment Pharmacol Ther* 2008 Feb 01;27(3):274-282 [FREE Full text] [doi: [10.1111/j.1365-2036.2007.03572.x](https://doi.org/10.1111/j.1365-2036.2007.03572.x)] [Medline: [17996017](https://pubmed.ncbi.nlm.nih.gov/17996017/)]
15. Turner CW, Lewis JR, Nielsen J. Determining usability test sample size internet. In: *International Encyclopedia of Ergonomics and Human Factors*. Boca Raton: CRC Press; 2006.
16. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res* 2017 Mar;27(4):591-608. [doi: [10.1177/1049732316665344](https://doi.org/10.1177/1049732316665344)] [Medline: [27670770](https://pubmed.ncbi.nlm.nih.gov/27670770/)]
17. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009 Aug 07;4:50 [FREE Full text] [doi: [10.1186/1748-5908-4-50](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]
18. Fonteyn ME, Vettese M, Lancaster DR, Bauer-Wu S. Developing a codebook to guide content analysis of expressive writing transcripts. *Appl Nurs Res* 2008 Aug;21(3):165-168. [doi: [10.1016/j.apnr.2006.08.005](https://doi.org/10.1016/j.apnr.2006.08.005)] [Medline: [18684411](https://pubmed.ncbi.nlm.nih.gov/18684411/)]
19. Chenail R. Interviewing the investigator: strategies for addressing instrumentation and researcher bias concerns in qualitative research. *Qual Report* 2011;16(1):255-262 [FREE Full text] [doi: [10.46743/2160-3715/2011.1051](https://doi.org/10.46743/2160-3715/2011.1051)]
20. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Adm Policy Ment Health* 2015 Sep;42(5):533-544 [FREE Full text] [doi: [10.1007/s10488-013-0528-y](https://doi.org/10.1007/s10488-013-0528-y)] [Medline: [24193818](https://pubmed.ncbi.nlm.nih.gov/24193818/)]
21. Denzin N. *The Research Act: A Theoretical Introduction to Sociological Methods*. Oxfordshire: Routledge; 2017.
22. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan W, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med* 2017 Sep;15(5):419-426 [FREE Full text] [doi: [10.1370/afm.2121](https://doi.org/10.1370/afm.2121)] [Medline: [28893811](https://pubmed.ncbi.nlm.nih.gov/28893811/)]

23. Bayley K, Belnap T, Savitz L, Masica A, Shah N, Fleming N. Challenges in using electronic health record data for CER: experience of 4 learning organizations and solutions applied. *Med Care* 2013 Aug;51(8 Suppl 3):S80-S86. [doi: [10.1097/MLR.0b013e31829b1d48](https://doi.org/10.1097/MLR.0b013e31829b1d48)] [Medline: [23774512](https://pubmed.ncbi.nlm.nih.gov/23774512/)]
24. Norheim OF, Baltussen R, Johri M, Chisholm D, Nord E, Brock D, et al. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Eff Resour Alloc* 2014 Aug 29;12:18 [FREE Full text] [doi: [10.1186/1478-7547-12-18](https://doi.org/10.1186/1478-7547-12-18)] [Medline: [25246855](https://pubmed.ncbi.nlm.nih.gov/25246855/)]
25. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behav Res Methods Instrum Comput* 2003 Aug;35(3):379-383. [doi: [10.3758/bf03195514](https://doi.org/10.3758/bf03195514)] [Medline: [14587545](https://pubmed.ncbi.nlm.nih.gov/14587545/)]
26. Aiyegbusi OL. Key methodological considerations for usability testing of electronic patient-reported outcome (ePRO) systems. *Qual Life Res* 2020 Feb;29(2):325-333 [FREE Full text] [doi: [10.1007/s11136-019-02329-z](https://doi.org/10.1007/s11136-019-02329-z)] [Medline: [31691202](https://pubmed.ncbi.nlm.nih.gov/31691202/)]
27. Glaser B, Strauss A. *Discovery of Grounded Theory: Strategies for Qualitative Research*. New Jersey: Aldine Transaction; 2009.
28. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6:42 [FREE Full text] [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]

Abbreviations

CCTS: Cancer Care Tracking System

CFIR: Consolidated Framework for Implementation Research

EMR: electronic medical record

HCV: hepatitis C virus

P-CIMS: Population-based Cirrhosis Identification and Management System

VA: Department of Veterans Affairs

VHA: Veterans Health Administration

Edited by G Eysenbach; submitted 08.02.21; peer-reviewed by SM Ayyoubzadeh, X He, Y Chu, S Rostam Niakan Kalhori ; comments to author 29.03.21; revised version received 23.07.21; accepted 18.09.21; published 09.11.21.

Please cite as:

Javier SJ, Wu J, Smith DL, Kanwal F, Martin LA, Clark J, Midboe AM

A Web-Based, Population-Based Cirrhosis Identification and Management System for Improving Cirrhosis Care: Qualitative Formative Evaluation

JMIR Form Res 2021;5(11):e27748

URL: <https://formative.jmir.org/2021/11/e27748>

doi: [10.2196/27748](https://doi.org/10.2196/27748)

PMID: [34751653](https://pubmed.ncbi.nlm.nih.gov/34751653/)

©Sarah J Javier, Justina Wu, Donna L Smith, Fasiha Kanwal, Lindsey A Martin, Jack Clark, Amanda M Midboe. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 09.11.2021. 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

Validation of a Mobile Health Technology Platform (FeverTracker) for Malaria Surveillance in India: Development and Usability Study

Ipsita Pal Bhowmick^{1*}, PhD; Dibyajyoti Chutia^{2*}, PhD; Avinash Chouhan², MTech; Nilay Nishant², MTech; P L N Raju², MTech; Kanwar Narain¹, PhD; Harpreet Kaur³, PhD; Rocky Pebam², PhD; Jayanta Debnath¹, BSc; Rabindra Tripura¹, MSc; Kongkona Gogoi¹, MCA; Suman Ch Nag¹, BSc; Aatreyee Nath², MSc; Debabrata Tripathy¹, PhD; Jotish Debbarma¹, BMLT; Nirapada Das¹, BMLT; Ujjwal Sarkar¹, MSc; Rislyn Debbarma¹, MSc; Rajashree Roy¹, MSc; Bishal Debnath¹, BA, BEd; Dipanjan Dasgupta¹, DMLT; Suraj Debbarma¹, HSC; Kamal Joy Tripura¹, BA; Guneram Reang¹, BA; Amit Sharma^{4,5}, BSc, PhD; Manju Rahi³, MBBS, MD; Jyoti Chhibber-Goel⁵, PhD

¹Regional Medical Research Centre-Northeastern Region, Indian Council of Medical Research, Dibrugarh, India

²North Eastern Space Applications Centre, Umam, India

³Indian Council of Medical Research, Delhi, India

⁴National Institute of Malaria Research, Indian Council of Medical Research, Dwarka, Delhi, India

⁵International Centre for Genetic Engineering and Biotechnology, New Delhi, India

*these authors contributed equally

Corresponding Author:

Jyoti Chhibber-Goel, PhD

International Centre for Genetic Engineering and Biotechnology

Vasant Kunj

New Delhi, 110067

India

Phone: 91 011 26741358 ext 170

Email: jyotichhibbergoel@gmail.com

Abstract

Background: A surveillance system is the foundation for disease prevention and control. Malaria surveillance is crucial for tracking regional and temporal patterns in disease incidence, assisting in recorded details, timely reporting, and frequency of analysis.

Objective: In this study, we aim to develop an integrated surveillance graphical app called FeverTracker, which has been designed to assist the community and health care workers in digital surveillance and thereby contribute toward malaria control and elimination.

Methods: FeverTracker uses a geographic information system and is linked to a web app with automated data digitization, SMS text messaging, and advisory instructions, thereby allowing immediate notification of individual cases to district and state health authorities in real time.

Results: The use of FeverTracker for malaria surveillance is evident, given the archaic paper-based surveillance tools used currently. The use of the app in 19 tribal villages of the Dhalai district in Tripura, India, assisted in the surveillance of 1880 suspected malaria patients and confirmed malaria infection in 93.4% (114/122; *Plasmodium falciparum*), 4.9% (6/122; *P vivax*), and 1.6% (2/122; *P falciparum/P vivax* mixed infection) of cases. Digital tools such as FeverTracker will be critical in integrating disease surveillance, and they offer instant data digitization for downstream processing.

Conclusions: The use of this technology in health care and research will strengthen the ongoing efforts to eliminate malaria. Moreover, FeverTracker provides a modifiable template for deployment in other disease systems.

(JMIR Form Res 2021;5(11):e28951) doi:[10.2196/28951](https://doi.org/10.2196/28951)

KEYWORDS

fever; health system; mHealth app; malaria; surveillance; mobile phone

Introduction

Background

Surveillance is defined as “the continuous and systematic collection, analysis, and interpretation of disease-specific data, and the use of that data in the planning, implementation, and evaluation of public health practice” [1]. Surveillance of malaria cases assists the health ministry to (1) calculate disease burden, (2) monitor changing disease patterns, (3) design effective health interventions, and (4) evaluate the impact of malaria control programs, thereby curtailing transmission [2]. Nationally coordinated strategies intensify control efforts directed toward populations with the highest disease transmission by means of targeted intervention [2,3]. Nations with weak surveillance systems fail to assess disease trends and plan interventions, thereby resulting in continued transmission of malaria. The 2016-2030 Global Technological Strategy for Malaria of the World Health Organization (WHO) encourages national malaria surveillance programs to accelerate progress from malaria control to elimination [4,5]. Surveillance systems comprise the community, health workers, tools, and structures required to generate, analyze, and interpret data. Subsequently, decision-makers use the data for planning and targeting interventions. A combination of effective activities such as passive case detection (PCD; ie, febrile patients reach out to health facilities and get diagnosed) and active case detection (ACD; ie, patients with malaria are detected by health workers through active search in the community) are used by the surveillance systems [6]. Health care workers engage in malaria surveillance via ACD and PCD at all levels of the health system and communicate data from the field to the district and state health levels. This is the usual channel of data communication deployed in malaria-endemic regions for malaria control and elimination.

Despite established surveillance strategies, malaria-monitoring technologies (eg, rapid diagnostic tests [RDT], data capture on parasites and vector drug or insecticide resistance, and disease transmission markers) in many countries remain insufficient to support targets of malaria elimination [7-9]. The National Vector Borne Disease Control Programme (NVBDCP) of India is the largest surveillance network in the world [2]. Despite a well-defined active and passive surveillance system and a decreasing trend in malaria cases, gross discrepancies between reported malaria incidence in India, and the estimated malaria cases in the WHO published World Malaria Report have been noted [10,11]. A recent study estimated malaria incidence to be four-fold greater than one million reported by the NVBDCP but three-fold less than 13 million estimated by the WHO, whereas the estimated deaths were 93-fold more than the average 313 deaths reported by the national malaria program in 2015-2016 [12]. NVBDCP surveillance misses malaria cases diagnosed and treated by private and other health sectors; for instance, those operating in remote areas must be captured to determine the true malaria burden [13]. In addition to the significant gap between the reported and estimated malaria cases and deaths, in a country with approximately 1.3 billion people, delayed communication of the data from the grassroots-level health workers (especially in remote areas) to the district and

further up to the state level results in the lack of prompt data compilation and analysis and hence timely and appropriate mitigation action. The challenge lies in the switch from aggregated, paper-based reporting systems to near real-time, case-based electronic systems [14]. Furthermore, malaria transmission is not homogeneously distributed either geographically or within populations. The WHO emphasizes the need to identify malaria transmission foci and *hotspots* within endemic regions [1]. Failure to identify hotspots would result in sustained residual malaria transmission, thus undermining malaria control programs.

Objectives

The use of digital technology via the implementation of smartphone surveillance apps by the community for self-reporting and setting up digital information and contact chains at all levels of health care would be beneficial for effective malaria surveillance [1,13,14]. This avenue is equally applicable in urban and rural areas.

The worldwide smartphone internet traffic accounts for approximately 51% of the total global web-based traffic, wherein smartphone-based internet use is often a major way to access information worldwide [15]. Furthermore, mobile devices (excluding tablets) account for approximately 80% of the total internet traffic in India, with approximately 448 million smartphone users as of December 2020 [15]. Here, we report the establishment of a new smartphone mobile health app called *FeverTracker*. The app connects the information provided by the user to the nearest health care center, thereby allowing the health care staff to reach symptomatic patients and perform malaria RDT and slides as required. Validation of the app in 19 malaria-endemic tribal villages situated in remote, forested hilly regions in Dhalai district of Tripura state for a period of 20 months was performed. The app use resulted in a shortening of time between the collection of data and its reporting by a month. The app has served to monitor malaria in an integrated fashion, giving an all-in-one tool to the 19 health care workers involved in malaria surveillance. To date, the app has been used to screen 1880 suspected patients and was able to detect 122 malaria cases (114/122, 93.4% *Plasmodium falciparum*; 6/122, 4.9% *P vivax*; and 2/122, 1.6% *P falciparum/P vivax* mixed infection cases). The use of the FeverTracker app aims to strengthen malaria surveillance and provide instant digitization of epidemiological data.

Methods

App Development

FeverTracker is based on the concept of a progressive web app using the Ionic (Drifty) and Cordova (developed by Joe Bowser, Michael Brooks and Team) platforms. The app can capture ground information on malaria incidents attributed to geographic positional information, geotagged photos, and other relevant data as per the prescribed format. FeverTracker supports multilingual data to cater to local requirements and inclusive use. Detailed symptom information and use cases are included in the app. Advisory regarding drug dose information based on user symptoms is also included. The app currently supports Android OS 4.4 (Google) or higher. A hypertext preprocessor

(PHP; developed by Rasmus Lerdorf) webserver was used to deploy the app, and this app works on client-server architecture where in the backend, PHP and MySQL servers (Oracle Corporation) are used. The app also works in an offline mode, where the SQLite database (developed by D Richard Hipp) is used for record management. All communication between the mobile app and the backend server is secured using HTTP secure (European Organization for Nuclear Research). The mobile app is interlinked with an interactive web portal for data visualization, summaries, and data downloads. The app adheres to privacy laws and policies while seeking permission to access location and permission to use text messages, mobile data, and local storage. In addition to automated data capture, FeverTracker is linked to SMS text messaging and an advisory instruction system to notify the district or state response center according to national guidelines.

Portal Development

The FeverTracker portal is built on top of an open-source stack and comprises major functionalities such as on-the-fly analytics, integration with mobile apps, visualization of rich and interactive geographic information system (GIS) maps, information box, print maps with statistics, and data updates. The app comprises a 3-tier architecture: (1) database, (2) GIS server, and (3) web client mapping. The architectural design follows a service-oriented architecture, where tier 1 consists of a database layer hosting Postgres (developed by Michael Stonebraker and team) or the MySQL database, wherein all the location information along with attributes are stored. In tier 2, the GIS server in use is geo-server, and the reporting tool used is map fish print (OSGeo project). The geospatial data are served via a geo-server and follow the open geospatial consortium standards, where the data serve as web mapping services as well as web feature services. The client side uses leaflet (developed by Vladimir Agafonkin) as a map display tool and D3js-based customization (developed by Mike Bostock, Jason Davies, Jeffrey Heer, and Vadim Ogievetsky) for querying across the parameters. The apps are deployed with a leaflet in the frontend and backend as PHP.

Deployment of the App

The app was deployed in 19 villages of the Dhalai district of the state of Tripura in India, involving 19 health care workers that included 2 multipurpose workers (MPWs) posted at Gurudhan and Shikaribari subcenters and 3 Accredited Social Health Activists (ASHAs) who oversee 6 villages under these subcenters. In addition, 14 village volunteers designated for malaria surveillance in these 19 villages trained jointly by state and project were recruited for the deployment of the app. Fever

surveillance in the community (ACD) or cases reported at subcenters (PCD) as well as seasonal mass surveillance undertaken by the state in the area was performed using the app. The health care workers using the app were given mobile handsets and power banks as backups because of frequent power cuts in these remote villages. A demonstration and hands-on training were provided to all app users.

Survey for App Users

A survey using Google Forms was conducted among the 19 health care workers using the app for suspected malaria patient surveillance. The survey questions were in 2 local languages (1) Bengali and (2) Kokborok. The answers were translated into English.

Preparation of Ecological Maps

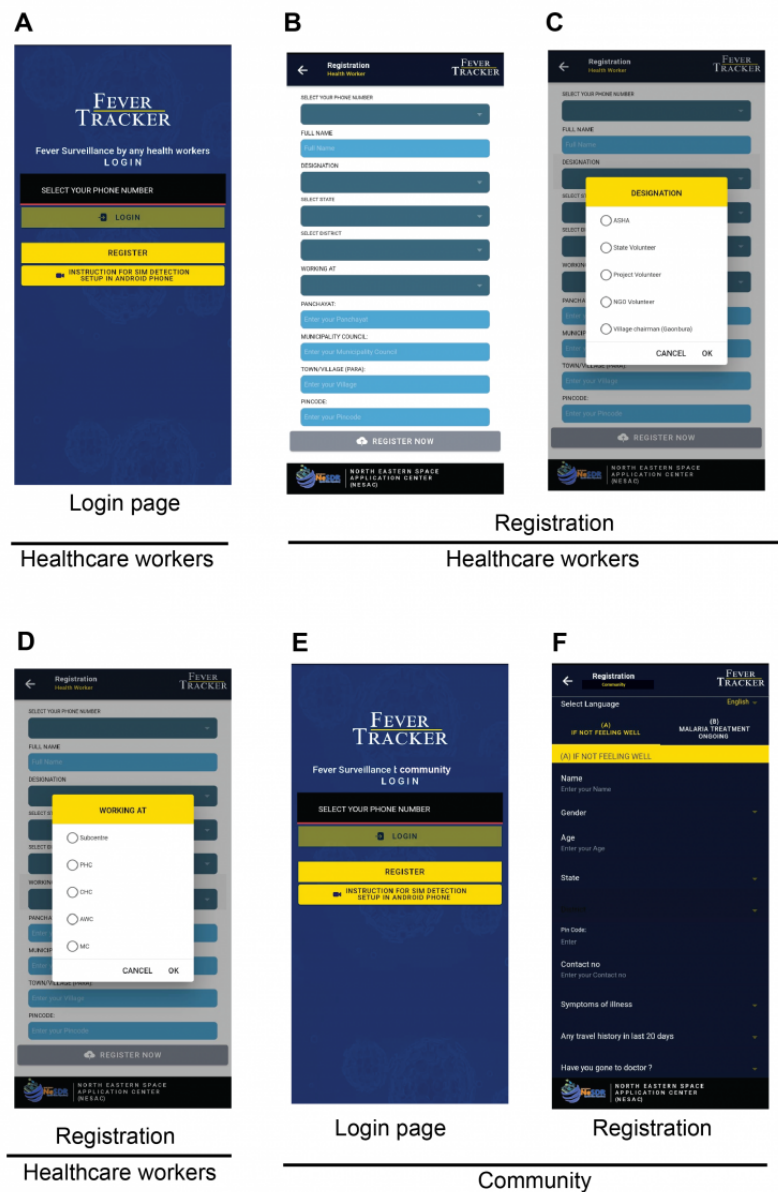
The land use land cover mapping was prepared using the orthorectified Indian remote sensing satellite data, Cartosat-1 (2.5 m) and Linear Imaging Self Scanning Sensor-IV (5.8 m). On-screen visual interpretation techniques were used on a GIS platform. Major land use and land cover categories along with subcategories were delineated and updated using the latest data (2019) on the spatial layer. The spatial layer was prepared under National Remote Sensing Centre/Indian Space Research Organization's space-based information support for decentralized planning at the panchayat level. Mapping was performed at 1:10,000 scales. The geolocations obtained from the FeverTracker app from villages screened under the Gurudhan and Shikaribari subcenters of Ambassa and Ganganagar were plotted on these ecological maps.

Results

Multifaceted Surveillance

The registered health care workers involved in door-to-door screening (ACD protocol) for malaria surveillance can use this app (Figure 1A). The app allows health care workers to register via their phone number and create a unique ID for future log-ins (Figure 1B). The list of health care workers involved is as follows: ASHAs, auxiliary nurse midwife, MPWs, community health officers at the field level and village volunteers, laboratory technicians, and medical officers engaged in testing and interpreting the results (Figure 1C). Notably, the app has the provision for private practitioners, laboratories, pharmaceuticals, and security forces to enter the data. The different levels of data collection and transmission will work for fever screening in the villages and at the subcenter, primary health centers, block level, district hospitals, or clinics (Figure 1D).

Figure 1. The log-in and registration page of the FeverTracker mobile health app. The community (A) log-in page and (B) registration page with required contact details, symptoms, travel history, and physician visit is shown. (C-F) The app allows the health care workers to log-in and register with their credentials and register, allowing them to perform surveillance at the grassroots level. Names of the persons and the villages have been removed to retain privacy.



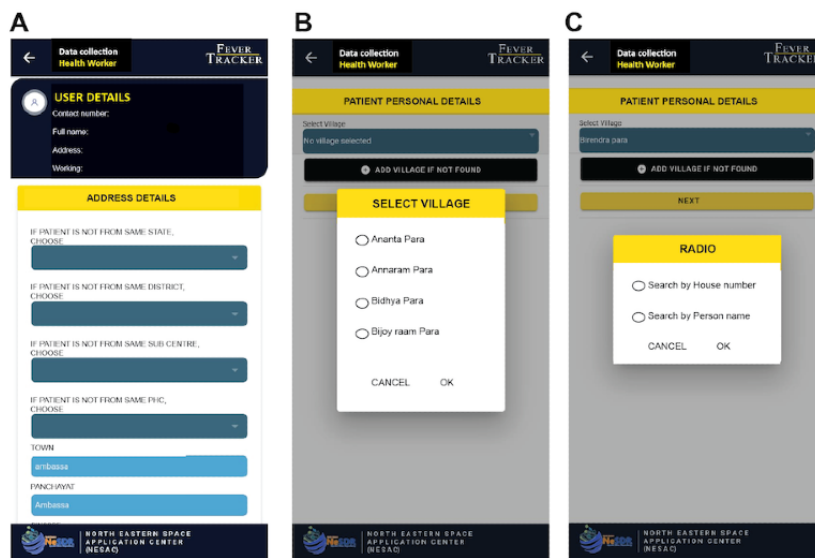
In addition, the app has provision for the participation of community via self-assessment and then to report malaria symptoms to health care workers, which will facilitate prompt health seeking and thus improve surveillance and malaria control. FeverTracker has features to assist the community in understanding their fever-related symptoms and to report them to health authorities [3]. This will supplement PCD with people self-reporting through the app and hence enhance surveillance with the help of at-risk communities. The app allows the patients screened for malaria via the private sector to be added as well as tracked for the accuracy of their treatment regimen, adherence, and compliance to the malaria drugs. In case of nonmalarial diseases, the individuals can be tested for other febrile illnesses. Further, based on their symptoms clustering, the outbreak of different diseases can be assessed and predicted.

FeverTracker allows users to enter their details and register via cell phone numbers that are stored in the health center database in a systematic and traceable format (Figure 1E and F). The app further connects the information provided by the user to the nearest malaria surveillance center. On receiving the information on a suspected malaria case in their surveillance region, health care workers can take appropriate action, for example, perform rapid diagnostics and slide microscopy to confirm malaria infection. In the case of a positive test result, health care workers are notified, and malaria treatment is provided. This key feature is further discussed and elaborated in later sections. In addition, if the user has been tested and treated, the app allows the user to report these details along with the complete course of treatment.

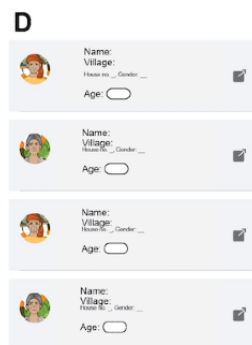
Thus, the app allows both the community and health care workers to integrate in ACD and PCD fever surveillance. Each can enlist their address and details, including name, age, address, phone number, and upload a photograph (optional; Figure 2A-C). The details of the household members filled under the personal and address detail columns are shown in Figure 2D. At the time of data collection, information on the symptoms of the patients is entered by the health care workers or by the community when self-reporting (Figure 3A). In addition, if multiple individuals from a single household were tested, the

details of each household member are entered (Figure 3B) along with the location of the testing center or place (Figure 3C). To ensure that this process is smooth, accurate, and not tedious for health care workers, pictorial options are provided to select a wide range of symptoms in malaria and other vector-borne diseases, as well as the RDT result outcome (Figure 3A and D) and the treatment regime to be followed by the patients who test positive for malaria according to the national guidelines (Figure 3E).

Figure 2. Data collection by health care workers using FeverTracker. (A-C) The contact details of the individual being surveyed or in case of a self-reported case are entered or verified by the health care workers. (D) The information on multiple family members of a single household is displayed on the app for the ease of health care workers during data collection and sampling. Names of the persons and the villages have been removed to retain privacy.

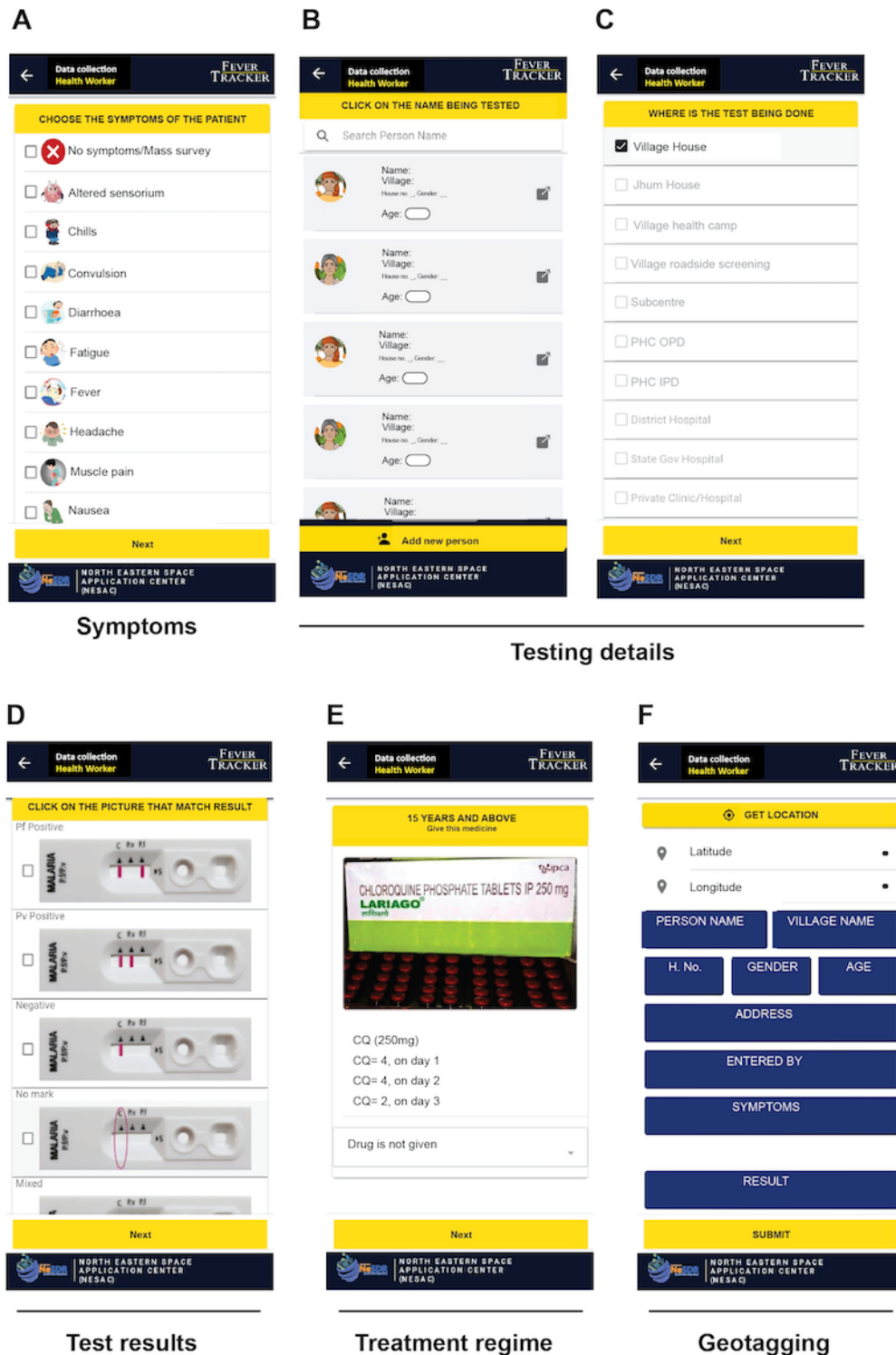


Patient details



Information on the members of a single household

Figure 3. The key features of the surveillance performed using the FeverTracker. (A) The health care workers and community members, on registration, can choose the disease symptoms. The testing details including (B) the name of the individual being tested, (C) the location of the test, (D) the test results, and (E) the medicine prescribed are shown. The app includes geolocation information to assist health authorities in data visualization. Names of the persons and the villages have been removed to retain privacy.



Offline Data Capture and Geomapping

For effective surveillance in remote parts of India with limited or no access to the mobile network, the FeverTracker app is functional offline. The app provides an offline data entry feature, wherein data remain stored on the device and are uploaded to the server when the network is available. Each data entry on

FeverTracker includes geolocation information (Figure 3F). Geomapping assists health authorities in visualizing district and state-wide data to locate malaria clusters and regions that require urgent attention. Furthermore, records of individuals who previously tested positive for malaria are available to state health officials to help assist and track the travel history of malaria cases across states within the country. Currently, the app is

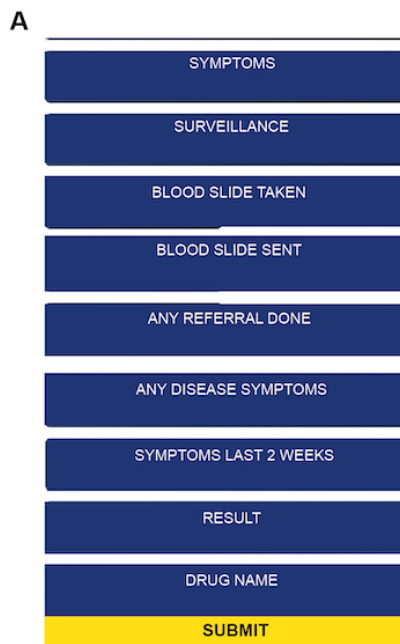
operational in 19 tribal villages in the Dhalai district of the state of Tripura, India.

Real-Time Data Capture and Coordinated Digitization of Surveillance

Digitization of data is an important and indispensable tool for surveillance programs to manage, visualize, and analyze data collected for any disease. The data capture constitutes multiple steps and involves health care workers from different strata, such as (1) a data entry operator or data manager who can enter the data obtained from the health workers, (2) data from primary health care centers that are sent to block level and then collated for district-level compilation and analysis, and (3) reports compiled or generated are sent from the districts to the state surveillance teams. The data generated via the above surveillance are the basis for the implementation of policies, strategies, and operational activities. Incomplete data or delays in data communication prolong malaria control action. In this scenario, our app allows the community to participate in malaria surveillance at the grassroots level. Health care workers receive an update on the information input by the community, which allows them to focus on individuals and areas requiring immediate attention. At the time of submission of the data collected via the FeverTracker app, the following information is available to the health authorities for further action: (1)

symptoms; (2) surveillance mode (ACD, PCD, or mass surveillance); (3) blood smear details; (4) duration of symptoms; (5) RDT test results, if any; and (6) medicine prescribed (Figure 4A). The surveillance team can edit and visualize the data on their mobile apps and on the web application dashboard (Figure 4B). After verification, the data entered into the app are communicated via an SMS text message to the concerned health care personnel and authorities. These data include test results with patient details and treatment status, thus allowing quick communication. The data from the FeverTracker main system are automatically digitized and made available via data sheets that can be used by the concerned district and state health workers and officials. Figure 4B shows the image of the data sheet with a patient entry highlighted along with the corresponding SMS text message and image of the patient with a date and time stamp to demonstrate the accuracy and usefulness of the app. The uploaded data can also be monitored by health officials via a web application platform in real time, allowing them to filter the data based on the set criteria for further analysis of malaria hotspots. In brief, the data input by the community and the data captured by health care workers are digitized in near real time and linked to the health centers via the SMS text message and advisory instructions system to notify the district or state response center.

Figure 4. The deployment of the FeverTracker. (A) The malaria surveillance data collected as visualized within the app is shown. (B) The data transmitted and stored in the server for visualization are shown. The data are digitized via the web platform in a presentable format with patient details, rapid diagnostic test results, symptoms, and drugs prescribed along with geolocations. The screenshot of the SMS text message received by the health care worker on verification of data is shown. One patient data entry is highlighted in the data spreadsheet along with the corresponding SMS text message and the patient image. RDT: rapid diagnostic test.



B

Time stamp	Patient details	RDT testing	Healthcare workers	Symptoms & Drugs prescribed	Location														
id	Entry date & time	patient picture	patient name	village name	house number	gender	age	test result	pregnancy status	designation of healthcare worker	symptoms	duration of symptoms	drugs to be administered	drugs given Yes/No	if no, the reason	status of drug given	remarks	latitude	longitude
1260	27/11/20	1606445128.jpg	H...	Tamarai		Female	2	PF +ve	None	Volunteer	Fever	10	Artemether 20mg - Lumefantrine 120mg	Yes	Drugs Given	This drug has been swallowed by patient	Fever test	23.8...	91.9...
1261	27/11/20					Female	9	-ve	None	ASHA	Fever	1	None	None	None	None	undefined	23.8...	91.8...
1262	27/11/20					Female				ASHA	No Symptoms/ Mass survey	0	None	None	None	None	undefined	23.8...	91.8...
1264	27/11/20					Female				ASHA	No Symptoms/ Mass survey	0	None	None	None	None	undefined	23.8...	91.8...
1265	28/11/20					Female				ASHA	Fever	2	Lumerax-80	Yes	Drugs Given	This drug has been swallowed by patient	undefined	23.8...	91.9...
1266	01/12/20					Female				Junteer	Fever	4	None	None	None	None	RDT TEST	23.8...	91.8...
1267	01/12/20					Female				ASHA	Altered Sensorium	2	Lumerax-80	Yes	Drugs Given	This drug has been swallowed by patient	undefined	25.6...	91.9...
1268	02/12/20					Female				W/CHO at bcentre	Fever	2	Lumerax-40	Yes	Drugs Given	This drug has been swallowed by patient	Yes	23.9...	91.8...
1269	07/12/20					Female				Junteer	Fever	7	Artemether 20mg - Lumefantrine 120mg	Yes	Drugs Given	This drug has been swallowed by patient		23.7...	91.9...
1270	08/12/20					Female				ASHA	Fever	1	None	None	None	None	undefined	23.8...	91.8...
1271	08/12/20					Female	7	-ve	None	ASHA	Fever	2	None	None	None	None	undefined	23.8...	91.8...
1272	10/12/20	1607571535icmr.jpg				Male	6	-ve	None	Volunteer	Fever	2	None	None	None	None	Fever	23.7...	91.9...
1273	10/12/20	1607571936icmr.jpg				Female	9	-ve	None	Volunteer	Fever	3	None	None	None	None	Fever	23.7...	91.9...
1274	10/12/20	1607572263icmr.jpg				Female	7	Negative	None	Volunteer	Fever	2	None	None	None	None	Fever	23.7...	91.9...

PATIENT DETAILS
 Name:
 Age:
 Gender:
 Village:
 House no.:

DRUG DETAILS
 Drug name
 Yes or no
 Reasons, if no:

Backend Data Capture and Display

The mobile app-linked web application is equipped with analytical and mapping tools based on the data received from the app. The landing page has the following features for semiautomated live data streaming: (1) daily confirmed and cumulative cases spanning the current, recovered, and malaria deaths across the surveyed regions; (2) a GIS map based on the data captured via the FeverTracker app to identify high caseload zones showing malaria clusters and regions; (3) a map showing the malaria incidence for various districts; and (4) statistical

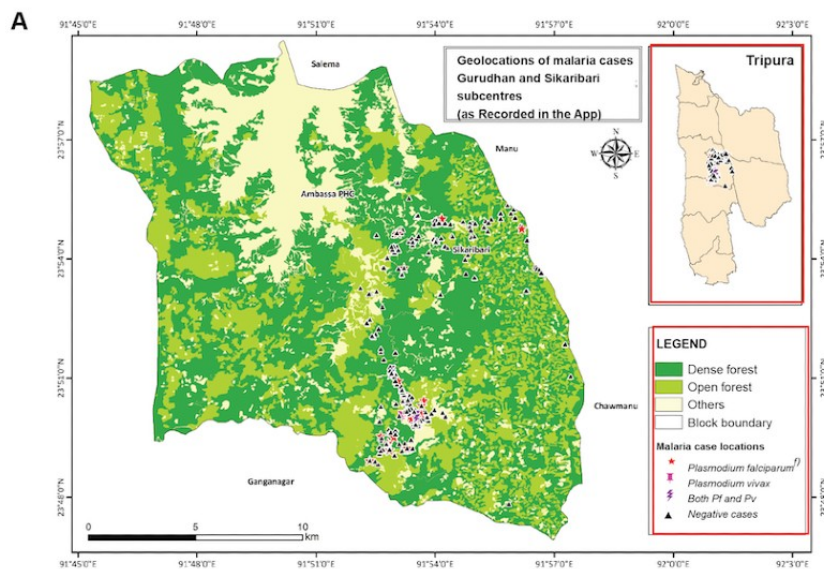
analysis of the samples tested for malaria along with age- and gender-wise information. As a proof of concept, the data collected from the app are shown in Figure 4. The information for each patient is visible on the app, as shown in Figure 4A. This information is linked to the web application to be viewed in an Excel format as raw data (Figure 4B). The additional feature in the FeverTracker app linking it to the web application allows it to be equipped with analytical and mapping tools based on the data received.

Data Collection on App Use to Understand Users' Perspective

The app was deployed in 19 malaria-endemic villages under Gurudhan, Shikaribari, Karnamani, and Maldapara subcenters of Ambassa and Ganganagar Primary Health Centre in Dhalai district of Tripura, India, for 20 months. These 19 health care workers (MPWs, ASHAs, and village volunteers) used the app to survey and screen a population of 4870 and perform RDT on 1880 people with fever suspected of malaria. Among these patients, 122 tested positive for malaria infection (114/122,

93.4% *P. falciparum* mono-infection; 6/122, 4.9% *P. vivax* mono-infection; and 2/122, 1.6% *P. falciparum/P. vivax* mixed infection cases), and the treatment provided was recorded. As part of this study, the health care workers involved in testing the app collated information on the symptoms, testing, treatment regime, and geolocations of suspected and positive patients with malaria. Figure 5A shows the geolocations plotted in the ecological maps of the state of Tripura, India. The area where the app was deployed was rural and a densely forested region where malaria surveillance was remotely monitored and tracked in real time.

Figure 5. Population surveyed and the survey for the health care workers. (A) The geolocations of the suspected malaria cases tested in Gurudhan and Shikaribari subcenter villages under Ambassa block in the district Dhalai, Tripura, obtained through the FeverTracker app are shown. The surveyed patient data were plotted in the land cover mapping of Ambassa block that was prepared using the orthorectified Indian remote sensing satellite data, Cartosat-1 (2.5 m) and Linear Imaging Self Scanning Sensor-IV (5.8 m), using on-screen visual interpretation techniques in the geographic information system platform. The mapping was done at 1:10,000 scales. The map of the state of Tripura with the district boundaries is shown in the inset. Data on the malaria cases from the mentioned area as registered via the app are marked (negative cases [blue], *Pf* [red], *Pv* [pink], and *Pf/Pv* [purple]). (B) The details of the survey conducted are given. The questions asked and the replies from the 19 health care workers obtained with the number of responses (in bracket) are listed.



B

S.No.	Questions	Responses (n=19)	
1	Is the app easy to use	Easy (19)	Difficult (0)
2	What is the frequency of your app use?	Always (10)	Sometimes (9)
3	Do you find the app better than paper based reporting?	Yes (18)	No (1)
4	Should the paper-based system be replaced by the app?	Yes (19)	No (0)
5	How much approx time it takes to use the app for each entry?	2-3 minutes (11)	4-5 minutes (8)
6	How easy or difficult was it to learn the app?	Easy (19)	Difficult (0)
7	How long did it take to learn the app?	1 day (13)	>2 days (6)
8	Do you find the Kokborok language option useful?	Yes (18)	No (1)
9	Do you find the picture based questions useful?	Yes (18)	No (1)
10	Do you find the household database where you can select from useful?	Yes (19)	No (0)
11	Would you like to continue using the app?	Yes (19)	No (0)
12	Do you find the drug guidance given in the app useful?	Yes (18)	No (1)

To assess the functionality of the app and its preference over a paper-based system, a survey was conducted among 19 health care workers using the app via Google Forms in English and 2 local languages (Bengali and Kokborok). The survey showed that all the users found the app easy to use and learn, with an average of 2 to 3 minutes per entry, and the users preferred to

continue using the app over the paper-based system (Figure 5B). The MPWs emphasized that SMS text messaging features of the app are most useful as SMS text messages are received when an ASHA or a village volunteer reports a malaria patient during an active surveillance in the village, in addition to the status of the treatment administered. This prevents delays in the

actions to be taken by MPWs. Furthermore, the local language and pictorial options make it easy and lucrative to be used by the ASHA and village volunteers. During paper-based recording of malaria cases, the translation of surveillance data from the local language to English is difficult and often results in delayed reporting. The additional feature of the village household database ensures that the data are verified, and no typing or data entry errors occur during data collection and further processing from local languages to English.

The health care workers, ASHAs, and village volunteers further emphasized that the minimal typing feature and the availability of the app in the local language made it easy and lucrative over paper-based reporting. Furthermore, the pictorial drug information in the app ensured an accurate dosage of drugs administered to the patients. The offline data capture feature of the app will be useful in remote areas of certain parts of India with no or poor connectivity. The use of the app has reduced the time delay from the collection of data to their reporting to the district or state health care system by over a month owing to the instantaneous SMS text message delivery system, which has made all concerned people in health systems aware of any positive case detection in the study area. The app was useful in these 19 remote villages with long electricity power cut sessions, which were overcome by supplying power banks for charging the mobile handsets.

Discussion

Principal Findings

The FeverTracker app was deployed in 19 malaria-endemic tribal villages situated in remote, poorly networked, hilly, forested regions of the state of Tripura in India via health care workers, ASHAs, and village volunteers. Over a period of 20 months, the use of the app resulted in regular data collection, shortening the time between data collection and processing by the health workers by over a month. According to a survey conducted for app users, FeverTracker was reportedly easy to use, beneficial for drug guidance, took less time to learn and use, reduced errors, and helped in village-level line listing with instant communication to the health system.

Strength and Limitations

The FeverTracker app has been particularly useful for malaria surveillance during the COVID-19 period with restricted mobility. The app records a range of fever-related symptoms, including COVID-19. This allowed the village volunteers and ASHAs to survey their respected villages, and the data collected reached the MPWs and district or state health officials without requiring the physical movement in or out of these villages. At the core of malaria management is extensive and coordinated surveillance of the community by health care workers at multiple levels. Among the challenges faced, limited mobility and resources along with the high number of target populations for surveillance compromises coverage and epidemiological data collection. Thus, the app offers a timely solution to these issues.

The app as a data collection tool, with minimal typing, pictorial, and automatic digitization features, reduces the time, effort, and resources that are presently dedicated toward paper-based

collection and digitization. This app also reduces the major current limitation of the mobility or physical travel of surveillance teams after the completion of the survey to provide updated data to the concerned health professionals. As a result, the entire process of data capture to analysis will be hastened, thus enabling rapid decision-making on the mitigation steps needed. FeverTracker has been designed such that it can be used by both private health care providers and the community to report malaria treatments undertaken at a private health care facility. Furthermore, the app will increase health seeking by the community as they would be self-reporting on suspicion of symptoms, preventing delays and other inconveniences associated with seeking the attention of a health care worker by shortening the channel of communication between the patient and the malaria health care personnel. These features of the FeverTracker app allow identifying and tracking of a vast number of patients outside the radar of malaria surveillance of the government and add robustness to the collection of epidemiological parameters required to track the progress of malaria control and elimination with the help of both the community and the health workers.

Conclusions

Currently, mobile health apps are being used to assist in monitoring patients outside hospitals, track vitals, or analyze medical images for physicians. More recently, the Food and Drug Administration started to approve health apps and devices [16]. Furthermore, the COVID-19 pandemic has shown the world utility of app-based data collection systems that are useful not only in prompt reporting, and hence data-based decisions, but also in creating awareness and alertness among communities [17,18]. In this light, FeverTracker has the potential to provide holistic support in the management of malaria control. It can also be used for joint surveillance and control programs of other eliminable vector-borne diseases, such as visceral leishmaniasis and lymphatic filariasis, which is desirable for the country at this juncture [19]. In sum, this app will allow (1) the omission of paper-based data capture (symptoms, test results, and medicines prescribed), (2) real-time communication of data to district and state surveillance teams, (3) pictorial representation and local language support for ground-level health workers, (4) offline data capture, (5) geotagging, and (6) link-up with web application-based dashboards that will allow the health authorities to locate transmission foci and hotspots. The app has been specifically designed, so it can be tailored according to the requirements of the government programs of any state or country. A gradual switch to app-based malaria surveillance will improve coverage, case data integration, data analysis, and epidemiological visualization and will inform malaria control programs in near real time. This robust *FeverTracker* app can be integrated within the health systems at the district, state, or national level to build a health care system network and can also be used for tracking acute fever or febrile illness other than malaria, such as dengue, leptospirosis, influenza, influenza A, rickettsial infections, Japanese encephalitis, and chikungunya and other infectious diseases, including COVID-19, while providing a long-term support to the health care system and public health research in the country.

Acknowledgments

The authors acknowledge the cooperation and support of the National Vector Borne Disease Control Programme and Department of Health and Family Welfare, officials of the state of Tripura, and health workers, as well as the village facilitators in Dhalai district, Tripura, for their help in the deployment study. This work was funded by the Indian Council of Medical Research extramural grant. AS was supported by the JC Bose fellowship. JCG was supported by DBT under the BioCARE scheme. IPB is also credited as a Corresponding Author.

Authors' Contributions

The malaria app was conceived by IPB and AS. DC, AC, NN, and PR developed the app and dashboard. KN, HK, JD, RT, KG, SCN, DT, JD, ND, US, RD, RR, BD, DD, SD, KT, and GR assisted in app deployment and testing in the northeastern state of Tripura in India. RP and AN constructed ecological plots using geolocation. IPB, AS, MR, and JCG coordinated the study and work. IPB and JCG wrote the manuscript. IPB, AS, MR, and JCG finalized the manuscript. All authors contributed to the work and have read and approved the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. Malaria Surveillance, Monitoring and Evaluation: A Reference Manual. Geneva: World Health Organization; 2018:1-208.
2. National Vector Borne Disease Control Programme. URL: <https://nvbdcp.gov.in> [accessed 2021-01-06]
3. Guidelines for diagnosis and treatment of malaria in India. National Institute of Malaria Research and Vector Borne Disease Control Programme. 2011. URL: https://nvbdcp.gov.in/Doc/Guidelines_for_Diagnosis_Treatment.pdf [accessed 2021-01-06]
4. World Health Organization. A Framework for Malaria Elimination. Geneva: World Health Organization; 2017:1-100.
5. World Health Organization. Global technical strategy for malaria 2016-2030, 2021 update. Geneva: World Health Organization; 2021:1-40.
6. Wickremasinghe R, Fernando SD, Thillekaratne J, Wijeyaratne PM, Wickremasinghe AR. Importance of active case detection in a malaria elimination programme. *Malar J* 2014 May 23;13(1):186-186 [FREE Full text] [doi: [10.1186/1475-2875-13-186](https://doi.org/10.1186/1475-2875-13-186)] [Medline: [24885972](https://pubmed.ncbi.nlm.nih.gov/24885972/)]
7. malERA Consultative Group on Health Systems and Operational Research. A research agenda for malaria eradication: health systems and operational research. *PLoS Med* 2011 Jan 25;8(1):e1000397 [FREE Full text] [doi: [10.1371/journal.pmed.1000397](https://doi.org/10.1371/journal.pmed.1000397)] [Medline: [21311588](https://pubmed.ncbi.nlm.nih.gov/21311588/)]
8. Kelly GC, Hale E, Donald W, Batarii W, Bugoro H, Nausien J, et al. A high-resolution geospatial surveillance-response system for malaria elimination in Solomon Islands and Vanuatu. *Malar J* 2013 Mar 21;12(1):108 [FREE Full text] [doi: [10.1186/1475-2875-12-108](https://doi.org/10.1186/1475-2875-12-108)] [Medline: [23514410](https://pubmed.ncbi.nlm.nih.gov/23514410/)]
9. Breman JG, Holloway C. Malaria surveillance counts. *Am J Trop Med Hyg* 2007 Dec;77(6 Suppl):36-47. [Medline: [18165473](https://pubmed.ncbi.nlm.nih.gov/18165473/)]
10. Rajvanshi H, Jain Y, Kaintura N, Soni C, Chandramohan R, Srinivasan R, et al. A comprehensive mobile application tool for disease surveillance, workforce management and supply chain management for Malaria Elimination Demonstration Project. *Malar J* 2021 Feb 16;20(1):91-99 [FREE Full text] [doi: [10.1186/s12936-021-03623-3](https://doi.org/10.1186/s12936-021-03623-3)] [Medline: [33593359](https://pubmed.ncbi.nlm.nih.gov/33593359/)]
11. World Health Organization. World Malaria Report 2020. Geneva: World Health Organization; 2020:1-299.
12. Kumar A, Chaturvedi HK, Mohanty AK, Sharma SK, Malhotra MS, Pandey A. Surveillance based estimation of burden of malaria in India, 2015-2016. *Malar J* 2020 Apr 16;19(1):156 [FREE Full text] [doi: [10.1186/s12936-020-03223-7](https://doi.org/10.1186/s12936-020-03223-7)] [Medline: [32299426](https://pubmed.ncbi.nlm.nih.gov/32299426/)]
13. Rahi M, Sharma A. For malaria elimination India needs a platform for data integration. *BMJ Glob Health* 2020 Dec 30;5(12):e004198 [FREE Full text] [doi: [10.1136/bmjgh-2020-004198](https://doi.org/10.1136/bmjgh-2020-004198)] [Medline: [33380414](https://pubmed.ncbi.nlm.nih.gov/33380414/)]
14. Rahi M, Das P, Sharma A. Malaria elimination in India requires additional surveillance mechanisms. *J Public Health (Oxf)* 2021 Apr 05(Online ahead of print):fdab106. [doi: [10.1093/pubmed/fdab106](https://doi.org/10.1093/pubmed/fdab106)] [Medline: [33823023](https://pubmed.ncbi.nlm.nih.gov/33823023/)]
15. Department of Telecommunications. Ministry of Communications, India. 2021. URL: <https://dot.gov.in> [accessed 2021-01-08]
16. Policy for device software functions and mobile medical applications - guidance for industry and Food and Drug Administration staff. US Food & Drug Administration. 2013. URL: <https://www.fda.gov/media/80958/download> [accessed 2021-01-06]
17. Rahi M, Das P, Sharma A. COVID-19 Mitigation steps provide a blueprint for malaria control and elimination. *Am J Trop Med Hyg* 2020 Jul;103(1):28-30 [FREE Full text] [doi: [10.4269/ajtmh.20-0394](https://doi.org/10.4269/ajtmh.20-0394)] [Medline: [32383431](https://pubmed.ncbi.nlm.nih.gov/32383431/)]

18. Rahi M, Baharia R, Das P, Chhibber-Goel J, Sharma A. Overlaying COVID-19 mitigation plans on malaria control infrastructures. *Trans R Soc Trop Med Hyg* 2021 Jan 07;115(1):6-8 [FREE Full text] [doi: [10.1093/trstmh/traa108](https://doi.org/10.1093/trstmh/traa108)] [Medline: [33045049](https://pubmed.ncbi.nlm.nih.gov/33045049/)]
19. Rahi M, Chaturvedi R, Das P, Sharma A. India can consider integration of three eliminable disease control programmes on malaria, lymphatic filariasis, and visceral leishmaniasis. *PLoS Pathog* 2021 May 20;17(5):e1009492 [FREE Full text] [doi: [10.1371/journal.ppat.1009492](https://doi.org/10.1371/journal.ppat.1009492)] [Medline: [34015028](https://pubmed.ncbi.nlm.nih.gov/34015028/)]

Abbreviations

ASHA: Accredited Social Health Activists
ACD: active case detection
GIS: geographic information system
MPW: multipurpose worker
NVBDCP: National Vector Borne Disease Control Programme
PCD: passive case detection
PHP: hypertext preprocessor
RDT: rapid diagnostic test
WHO: World Health Organization

Edited by G Eysenbach; submitted 22.03.21; peer-reviewed by W Lambert, F Holl, G Kumar; comments to author 02.06.21; revised version received 25.06.21; accepted 26.07.21; published 10.11.21.

Please cite as:

Pal Bhowmick I, Chutia D, Chouhan A, Nishant N, Raju PLN, Narain K, Kaur H, Pebam R, Debnath J, Tripura R, Gogoi K, Ch Nag S, Nath A, Tripathy D, Debbarma J, Das N, Sarkar U, Debbarma R, Roy R, Debnath B, Dasgupta D, Debbarma S, Joy Tripura K, Reang G, Sharma A, Rahi M, Chhibber-Goel J

Validation of a Mobile Health Technology Platform (FeverTracker) for Malaria Surveillance in India: Development and Usability Study

JMIR Form Res 2021;5(11):e28951

URL: <https://formative.jmir.org/2021/11/e28951>

doi: [10.2196/28951](https://doi.org/10.2196/28951)

PMID: [34757321](https://pubmed.ncbi.nlm.nih.gov/34757321/)

©Ipsita Pal Bhowmick, Dibyajyoti Chutia, Avinash Chouhan, Nilay Nishant, P L N Raju, Kanwar Narain, Harpreet Kaur, Rocky Pebam, Jayanta Debnath, Rabindra Tripura, Kongkona Gogoi, Suman Ch Nag, Aatreyee Nath, Debabrata Tripathy, Jotish Debbarma, Nirapada Das, Ujjwal Sarkar, Rislyn Debbarma, Rajashree Roy, Bishal Debnath, Dipanjan Dasgupta, Suraj Debbarma, Kamal Joy Tripura, Guneram Reang, Amit Sharma, Manju Rahi, Jyoti Chhibber-Goel. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.11.2021. 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 a Severity Score and Comparison With Validated Measures for Depression and Anxiety: Validation Study

William Lynch^{1*}, BA; Michael L Platt^{2,3,4*}, PhD; Adam Pardes^{1*}, MSE

¹NeuroFlow, Inc, Philadelphia, PA, United States

²Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

³Department of Psychology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, United States

⁴Marketing Department, The Wharton School of Business, University of Pennsylvania, Philadelphia, PA, United States

*all authors contributed equally

Corresponding Author:

Adam Pardes, MSE

NeuroFlow, Inc

111 S Independence Mall E

Suite 701

Philadelphia, PA

United States

Phone: 1 267 671 7316

Email: adam@neuroflow.com

Abstract

Background: Less than 10% of the individuals seeking behavioral health care receive measurement-based care (MBC). Technology has the potential to implement MBC in a secure and efficient manner. To test this idea, a mobile health (mHealth) platform was developed with the goal of making MBC easier to deliver by clinicians and more accessible to patients within integrated behavioral health care. Data from over 3000 users of the mHealth platform were used to develop an output severity score, a robust screening measure for depression and anxiety.

Objective: The aim of this study is to compare severity scores with scores from validated assessments for depression and anxiety and scores from clinician review to evaluate the potential added value of this new measure.

Methods: The severity score uses patient-reported and passively collected data related to behavioral health on an mHealth platform. An artificial intelligence-derived algorithm was developed that condenses behavioral health data into a single, quantifiable measure for longitudinal tracking of an individual's depression and anxiety symptoms. Linear regression and Bland-Altman analyses were used to evaluate the relationships and differences between severity scores and Personal Health Questionnaire-9 (PHQ-9) or Generalized Anxiety Disorder-7 (GAD-7) scores from over 35,000 mHealth platform users. The severity score was also compared with a review by a panel of expert clinicians for a subset of 250 individuals.

Results: Linear regression results showed a strong correlation between the severity score and PHQ-9 ($r=0.74$; $P<.001$) and GAD-7 ($r=0.80$; $P<.001$) changes. A strong positive correlation was also found between the severity score and expert panel clinical review ($r=0.80-0.84$; $P<.001$). However, Bland-Altman analysis and the evaluation of outliers on regression analysis showed that the severity score was significantly different from the PHQ-9.

Conclusions: Clinicians can reliably use the mHealth severity score as a proxy measure for screening and monitoring behavioral health symptoms longitudinally. The severity score may identify at-risk individuals who are not identified by the PHQ-9. Further research is warranted to evaluate the sensitivity and specificity of the severity score.

(*JMIR Form Res* 2021;5(11):e30313) doi:[10.2196/30313](https://doi.org/10.2196/30313)

KEYWORDS

PHQ-9; GAD-7; depression assessment; anxiety assessment; measurement-based care; integrated behavioral health

Introduction

Integrated Measurement-Based Behavioral Health Care

Measurement-based care (MBC) can be broadly defined as the use of continuous monitoring of patient data to inform and, as needed, redirect clinical care [1]. Behavioral health focuses on how behaviors impact a person's health both physically and mentally. Within the behavioral health field, MBC is considered an evidence-based practice, which is defined as relying on the use of data from controlled scientific studies reported in the published literature to make reasonable and conscientious decisions about clinical care [2]. As a framework for behavioral health care, MBC uses validated clinical assessments to measure symptoms of anxiety and depression before or during a patient-clinician interaction, and clinicians use these measurements to guide further clinical interventions [1]. MBC often occurs during cognitive behavioral therapy (CBT) or other psychological treatments [1,2]. Despite strong evidence that MBC improves behavioral health outcomes [1,3-6], less than 20% of behavioral health care providers use MBC [5,6]. Research has shown that barriers to the use of MBC include, but are not limited to, the costs in both time and resources to take measurements in addition to concerns about potential violations of confidentiality with paper and pen measurements [4,5].

Integrating behavioral health care into primary care has been proposed as a means of improving access to behavioral health care and is supported by both the American Association of Family Physicians and the American Psychiatric Association [7-10]. Globally, the integrated behavioral health model has demonstrated success in the United Kingdom with the web-based Beating the Blues program [11]. The program uses CBT and has been shown to be evidence-based and cost-effective in treating anxiety and depression by the primary care provider [11]. In the integrated behavioral health model, the primary care clinician screens patients for symptoms of mental health disorders at every visit and recommends and prescribes further assessment and treatment as needed, whether that be with medication, CBT, or other psychological treatments [9,10]. Patients are followed up by a care manager, and psychiatrists are available to consult with the primary care clinician as needed for any case. The integrated behavioral health model could help address the shortage of behavioral health care providers and low levels of access to behavioral health care, as approximately 75% of people in the United States have a primary care provider [12]. Globally, this percentage rises to 85% of people with primary care providers [13], demonstrating the imperative need to use primary care providers for behavioral health. Integrated behavioral health may also help address the unequal distribution of behavioral health care providers, as this model allows for remote screening and consultations. However, the integrated behavioral health model is not yet widely used, as only 4.3% of the primary care visits in 2016 included screening for behavioral health symptoms [14].

Given such low levels of depression or anxiety screening during primary care visits, it is perhaps not surprising that data are limited regarding the use of MBC within integrated behavioral health care. However, it is known that time, resources, and confidentiality are obstacles to the use of MBC by behavioral health providers who typically have 45 minutes with a patient weekly or biweekly [4,5]. There is also a paucity of data regarding the amount of time primary care physicians spend with individual patients. Typically, visits are scheduled every 15 minutes, and data suggest that the average visit may last 17-21 minutes [15,16], although a time-and-motion study suggests that only 53% of that time (9-11 minutes) is spent engaging with the patient directly [17]. Although current validated measures of depression and anxiety—the Personal Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorders-7 (GAD-7)—are brief and take ≤5 minutes to administer, providing MBC might take away 20% to 60% of the limited face-to-face time, during which many other clinical tasks must also be performed.

Unmet Behavioral Health Care Needs

Whether because of access to care, short time spent with physicians, or shortage of behavioral health clinicians, many individuals with depression and anxiety do not receive MBC, despite evidence that it improves outcomes [1]. Moreover, considering the epidemic proportions of depression and anxiety and related disability and mortality, there is an urgent need to find novel ways for primary care clinicians to provide comprehensive MBC in an integrated behavioral health care model. In addition, although the PHQ-9 and GAD-7 function efficiently on their own to identify at-risk patients, they alone may not capture the full scope of potential indicators of an at-risk patient beyond questions asked in the screening assessment.

Therefore, a mobile health (mHealth) platform was built with the goal of improving access to MBC at scale in an integrated behavioral health care model. There are some common concerns regarding smartphone apps for behavioral health care. For example, digital measurements are not necessarily evidence-based, may not provide treatment that is equivalent to that received from a trained clinician, and often have a low frequency of use and engagement by patients. Often, the measures used in apps were originally validated as paper and pen or pencil instruments, and it is not clear if digital versions have the same sensitivity and specificity. The number of available behavioral health apps for smartphones is in the tens of thousands [18], whereas a search on PubMed for clinical studies of such apps returns results in the thousands. The apps also deliver a wide variety of measures and interventions, making comparisons among apps and between apps and human-delivered (in-person or via telehealth) clinical interventions difficult, if not impossible, without such systematic studies. Concerns about the persistence of any effects are related to the fact that the use of behavioral health apps, on average, falls off precipitously, with only 3.3% of downloaded behavioral health apps used for more than 30 days [19]. As part of an effort to continue simplifying and increasing the adoption of MBC by clinicians and patients, a severity score was developed within the abovementioned mHealth platform. The severity score is a

single composite, relative nondiagnostic measure of behavioral health that can be tracked over time.

To understand the role of the severity score in MBC, the aim of this study is to validate the output measure against established scores of behavioral health assessments. Herein, we report the data comparing the severity score with scores from the PHQ-9 assessment, the GAD-7 assessment, and expert clinician consensus.

Methods

The mHealth Platform

In this study, all measurements were taken on the NeuroFlow mHealth platform (NeuroFlow, Inc), where all measurements were completed by patients on their own devices. Through the NeuroFlow app or desktop interface, patients record, track, and report their mood, sleep quality, stress level, and complete PHQ-9 and GAD-7 assessments regularly. The app also provides engaging educational videos based on users' self-reported scores and symptom changes over time. Clinicians with access to the mHealth platform send individualized links to their patients, which allows the patient to download the mobile app on their smartphone or access the platform through a website.

Although the latter features are not necessarily interventional or measurement oriented, they serve the goal of maintaining user engagement and lasting behavior change over time by supporting users. All measurements were securely stored in the Health Insurance Portability and Accountability Act (HIPAA)-compliant database, and when in use, a clinician accessed only their own patients' information via a digital dashboard that integrates with the electronic medical record. This integration ensured that taking the measurements required for MBC does not further reduce the already small amount of time that clinicians have with patients. Instead of performing

screening assessments, they can immediately see which of their patients may require a behavioral health intervention. The reporting system in the mHealth platform alerts clinicians about the patients who are not progressing or who have declining behavioral health. The features of the mobile app include daily self-rating scales for stress, mood, sleep, and pain; mindfulness tools; general health education; and step tracking. Among all mHealth app use, 30-day user retention is 70% compared with a typical 30-day retention rate of 3.3% for behavioral health mobile apps [20].

Development of the Severity Score

To create the composite severity score, we developed a proprietary algorithm that uses the measures recorded in the mHealth platform NeuroFlow by patients. The data set included deidentified records of over 3000 individuals who used the mHealth platform after it was assigned to them by a clinician between 2018 and 2019. Weighted variables included total scores on the PHQ-9 (27-point scale) or GAD-7 (21-point scale) measures taken within the app, self-reported sleep quality measures (scale of 0-10), self-reported mood measures (scale of 0-10), active behavioral health treatment (yes or no), frequency of an individual using specific activities within the app (collected passively), and whether the individual had endorsed having suicidal ideation (yes or no). For each measure of the severity score, descriptive statistics were calculated to determine the distribution of variables; central tendencies were calculated through mean and median and the spread of variable values through range, SD, and variance. Variables were assigned positive or negative weights based on expert clinician input, and these variables were combined to produce a severity score measure, using a proprietary artificial intelligence algorithm. Severity score measures ranged from 1 to 5, with 1 indicating a low risk for common behavioral health conditions (eg, depression and anxiety) and 5 indicating a high risk for such conditions (Figure 1).

Figure 1. Overview of the process for developing and assessing the severity score. GAD-7: Generalized Anxiety Disorder-7; PHQ-9: Personal Health Questionnaire-9.



Comparisons With Validated Assessments

The PHQ-9 is a 9-item questionnaire used to screen for depression in medical settings. Each individual item is scored as a potential total score from 0 to 27. PHQ-9 scores of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe depression, respectively. Under the scope of MBC, patients are asked by their care provider to complete assessments every 2 to 4 weeks. The clinical purpose of these assessments is to help support clinicians in making a diagnosis, to quantify depression symptoms, and to monitor changes over time to determine if treatment is making a difference [21,22]. The GAD-7 is a 7-item questionnaire used to screen for anxiety in medical settings; individuals are asked how often they have experienced certain feelings in the last 2 weeks on a scale of 0-3. The total possible score is 21, and scores of 5, 10, and 15 are considered cut-offs for the presence of mild, moderate, and severe anxiety, respectively [23]. These are components of the measurements that patients complete in the mHealth platform such that for every severity score logged in the platform, there is a corresponding overall PHQ-9 or overall GAD-7 score for that individual at that point in time.

Using deidentified records, the severity score for a given individual was plotted against their GAD-7 ($n=31,260$) and PHQ-9 ($n=36,324$) scores. For each comparison, we used linear regression to fit the slope of the line to our data and used Pearson product-moment correlation coefficient (r) to quantify and summarize the direction and magnitude of the relationship between our variables. The Pearson product-moment correlation coefficient r can be any number between -1 and 1 . The sign of r corresponds to the direction of the relationship between the variables and the number corresponding to the magnitude of the relationship between the variables.

The severity score measure and PHQ-9 scores were also compared using the Bland-Altman analysis, which compares the differences between 2 measures versus the mean of the 2 measures [24]. This analysis evaluated whether the two measurement methods returned the same results and therefore could be used interchangeably [25]. Because PHQ-9 scores vary from 0 to 27 and severity score measures vary from 1 to 5, we first converted PHQ-9 scores to the clinically meaningful categories of 1 for no depression (scores 0-4 on PHQ-9), 2 for mild depression (scores 5-9 on PHQ-9), 3 for moderate depression (scores 10-14 on PHQ-9), 4 for moderately severe depression (scores 15-19 on PHQ-9), and 5 for severe depression (scores 20-27 on PHQ-9). The differences between the converted PHQ-9 score and the severity score were then plotted against the average of the converted PHQ-9 and the severity score. To create the change scores for the severity score, we took the absolute change between a severity score taken at time point t and the previous severity score taken at time point $t-1$ for a given participant. We then applied the same calculation to create the change scores for the PHQ-9 and the GAD-7.

Comparison With Clinical Reviews

A panel of 6 behavioral-health expert clinicians, including licensed psychologists, licensed clinical social workers, and 1 mental health psychiatric board-certified registered nurse, provided 2 clinical reviews for each of the 250 individuals based on 2 different blinded presentations of data from deidentified patient records. As shown in Table 1, the first data set provided only 30-day average PHQ-9 and GAD-7 scores recorded in the app over a 30-day period for these 250 individuals. The second data set provided measures included in the severity score in addition to the 30-day average and maximal PHQ-9 and GAD-7 scores, including the suicidal ideation score from the PHQ-9, measures of sleep and mood, whether the individual reported

working with a behavioral health specialist, and whether severe depression or anxiety had been present in the last 30-day period. The 2 data sets were randomized independently to ensure that the order of the records was not repeated. Clinicians were asked

to assign a clinical rating of symptom severity to each of the 250 individual records for both data sets using a scale of 1 (low to minimal) to 5 (severe) (Textbox 1).

Table 1. Variables included in data set 2 evaluated by expert panel.

Variable	Possible scores ^a	Description
avg_phq9_score	0-27	The average PHQ-9 ^b score
max_phq9_score	0-27	The highest PHQ-9 score recorded
q9_max	0-3	The highest score recorded for the ninth question of the PHQ-9, which screens for suicidal ideation as follows: "Over the past 2 weeks, how often have you been bothered by thoughts that you would be better off dead or of hurting yourself in some way?" (0=not at all, 1=several days, 2=more than half the days, and 3=nearly every day)
avg_gad7_score	0-21	The average GAD-7 ^c score
max_gad7_score	0-21	The highest GAD-7 score recorded
avg_sleep_rating	0 (best)-4 (worst)	The average sleep rating
sleep_count	0-30	The number of sleep ratings
avg_mood_rating	0 (best)-4 (worst)	The average mood rating
mood_count	0-30	The number of mood ratings
has_bh_specialist	0=no and 1=yes	This value denotes whether a patient self-reported having a behavioral health clinician
is_severe ^d	0=no and 1=yes	This value denotes whether a severe GAD-7 or PHQ-9 score was recorded in the 30-day period before this period

^aAll scores except *is_severe* are for the same 30-day period.

^bPHQ-9: Physical Health Questionnaire-9.

^cGAD-7: General Anxiety Disorders-7.

^d*is_severe* is gathered from the last 30-day period.

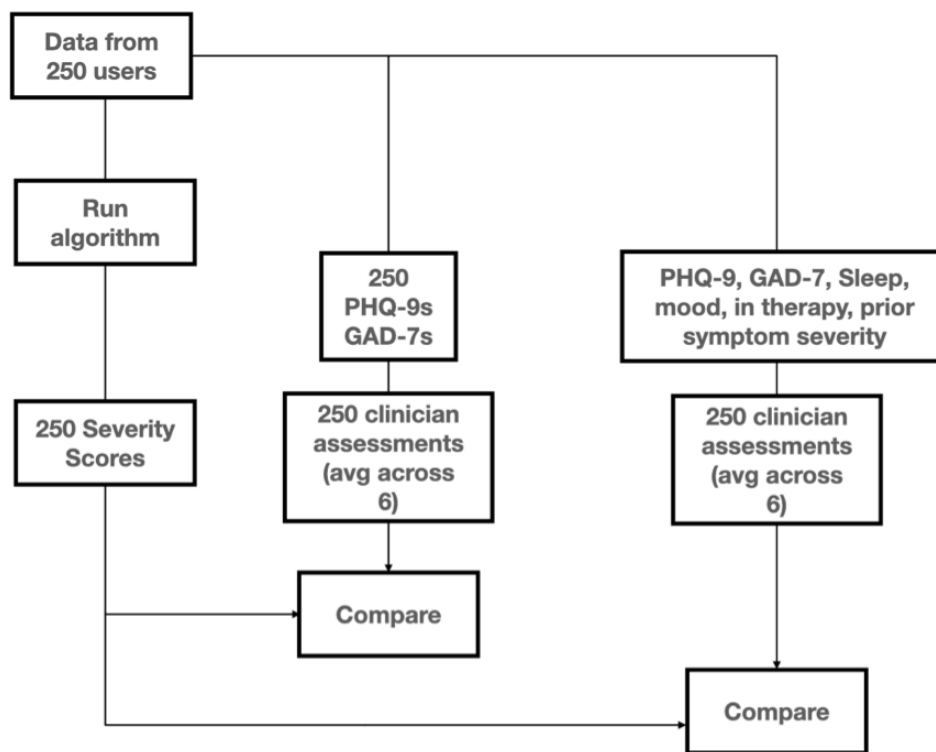
Textbox 1. Clinician symptom severity rating scale.

Scores
1: low to minimal severity
2: mild severity
3: moderate severity
4: moderately severe
5: severe

For each data set, the mean clinician expert score was plotted against the severity score for the same individual (n=250). Linear regression was calculated, and both the Pearson correlation coefficient and the Kendall tau measure were used

to evaluate the relationship between the mean clinician expert score and the severity score (Multimedia Appendix 1). Figure 2 provides an overview of how the severity score was compared with clinical reviews.

Figure 2. Overview of the process for comparing the severity score with clinicians’ reviews. avg: average; GAD-7: Generalized Anxiety Disorder-7; PHQ-9: Personal Health Questionnaire-9.



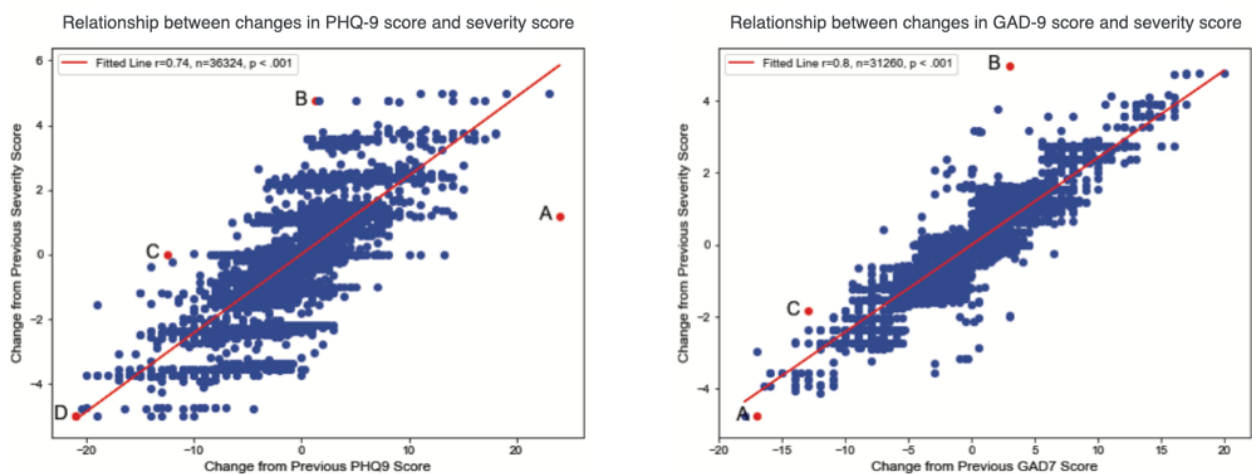
Results

Comparison of the Severity Score With Validated Assessments

Normal distribution of data points around the mean and median were found for all 3 changes measured (ie, severity score, PHQ-9, and GAD-7). This study reviewed the changes in scores to better understand the relationship between severity scores

and standardized assessments. The data for each measure were homoscedastic, supporting the use of linear regression for the analysis of possible correlations. After plotting the change in severity score versus the change in PHQ-9 scores for each individual record, we found a strong positive correlation between change in severity score and change in PHQ-9 score ($r=0.74$; $P<.001$; Figure 3). Similarly, changes in the severity score were strongly and positively correlated with changes in the GAD-7 score ($r=0.80$; $P<.001$; Figure 3).

Figure 3. Change in severity score versus change in Personal Health Questionnaire-9 (PHQ-9) score or Generalized Anxiety Disorders-7 (GAD-7) over a 30-day period.



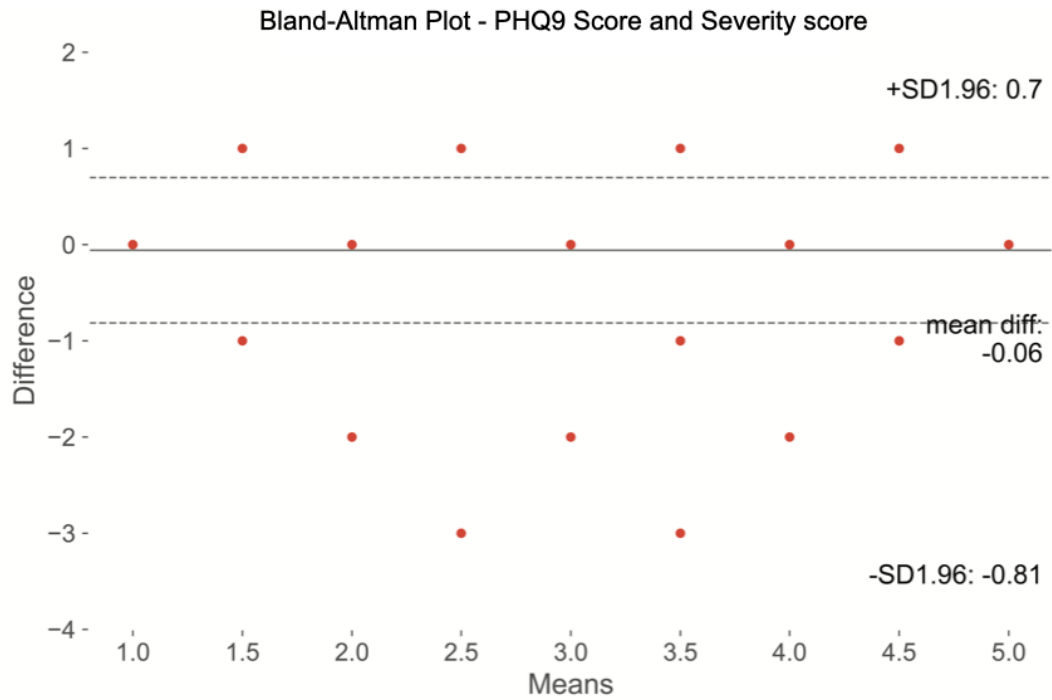
We also evaluated several data points that did not fit the linear regression and found that these were not errors or anomalies but rather were cases in which the severity score provided clinically meaningful information not captured by the PHQ-9

or GAD-7 score alone. For example, an individual whose PHQ-9 score changed from indicating minimal to moderately severe symptoms of depression (Figure 3) had a smaller change in the severity score because the severity score measure incorporated

their active involvement in mental health treatment and absence of suicidal ideation. In contrast, another individual had PHQ-9 score changes that reflected a decrease in symptoms of depression but did not have a drop in the severity score because the severity score incorporated ongoing thoughts of self-harm or suicidal ideation reported by this individual.

On the Bland-Altman analysis, comparing the PHQ-9 with the severity score, we found that the severity score measure was significantly different from the PHQ-9, with the difference between the two measures increasing around the mean of the two measures (Figure 4).

Figure 4. Bland-Altman analysis shows that the Personal Health Questionnaire-9 (PHQ-9) and the severity score are significantly different measures with differences increasing around the mean of the 2 measures.

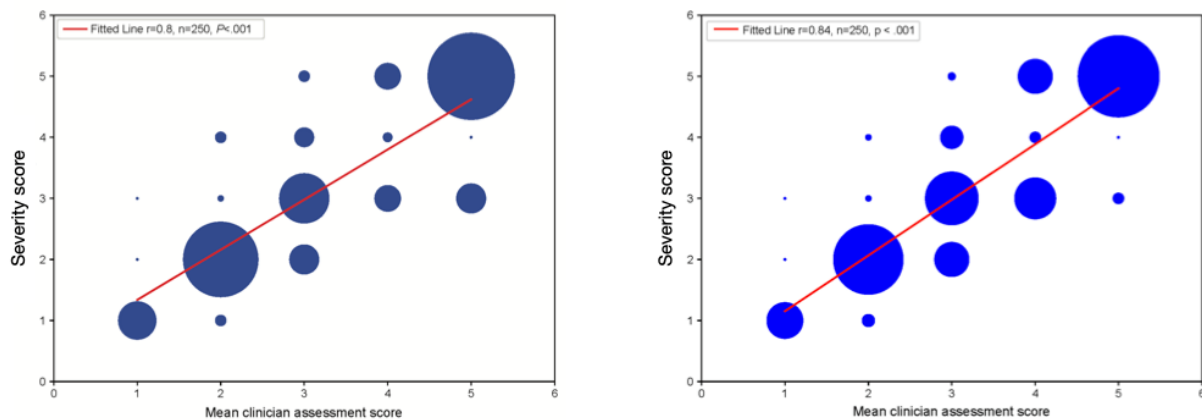


Comparison With Clinician Review

There were also strong correlations between the severity score and clinical experts’ reviews both when the expert clinical score was generated from only average PHQ-9 and GAD-7 scores ($r=0.80$; $P<.001$; Figure 5) and when the larger data set was used by the clinicians ($r=0.84$; $P<.001$; Figure 5). Linear

regression analysis showed a strong correlation between the severity score and clinician reviews for both data sets, with a slightly stronger correlation for the more comprehensive data set ($r=0.84$; $P<.001$) versus that for just the average PHQ-9 and GAD-7 scores ($r=.80$, $P<.001$). The 2 data sets came from the same 250 individuals and were separately randomized to ensure that the records were presented in different orders.

Figure 5. Severity score compared with mean clinical review score by expert clinician panel.



Discussion

Principal Findings

This study examined a new measure to be used in comprehensive MBC screening that considers validated assessments, subjective self-assessments, ongoing treatment, and app activity to gain full insight into the lives of patients. This study found relationships among the severity score, PHQ-9, and GAD-7. There was also a strong positive correlation between the severity score and expert panel clinical review, indicating that the severity score may serve as a relative measure of depression or anxiety. Findings from this study can inform further development of a comprehensive measure for clinicians to further identify at-risk individuals who may have otherwise been overlooked or not treated.

Integrating MBC into primary care has proven to be effective, but it is still not widely adopted [7]. For example, simply screening for the presence of depression or anxiety was performed in only 4.3% of primary care visits in 2016 [13]. Barriers to the implementation of MBC in primary care include systemic factors such as the lack of a consistently used measurement system and limited resources for training staff to use a measurement system in a HIPAA-compliant manner. In behavioral health care, HIPAA compliance is of particular importance because of the stigma surrounding behavioral health disorders that leads to discrimination and decreased quality of life [3-6,8,9,25]. At the individual patient level, concerns about confidentiality and the time burden for completing measurements are major barriers to the use of MBC. At the provider level, it is difficult to find time to measure behavioral health symptoms during primary care visits, given the high number of clinical tasks that must be prioritized and completed. Measurement-based care platforms and rating measures such as the severity score were developed with the goal of facilitating MBC integration into primary care by reducing time burden. For example, these types of platforms and rating scales take the measurement aspect outside of the clinical appointment and into the patients' daily activities with an engaging smartphone or desktop app.

Before investigating the feasibility and usability of the severity score in a clinical setting, it was crucial to first understand whether the severity score added value beyond the information provided by the PHQ-9 and the GAD-7, resulting in a more comprehensive understanding of the behavioral health profile of the user. The severity score is derived using an artificial intelligence-powered algorithm that uses both patient-reported measures taken outside the clinical visit and passively collected data through the use of an app. Regression analysis showed that the severity score correlated strongly with both the PHQ-9 and the GAD-7 scores with 0.88 sensitivity and specificity and 0.83 sensitivity and 0.84 specificity, respectively [21-23]. The severity score correlated strongly with clinician review based on PHQ-9 and GAD-7 scores alone ($r=0.80$; $P<.001$) and based on the analysis of a larger number of data reported using the mHealth platform ($r=0.84$; $P<.001$). These strong correlations suggest that the severity score can be used as a proxy measure for the presence of symptoms of depression and anxiety.

For example, if a clinician observed (on the clinician dashboard or electronic health record) that a patient had a severity score of 2 or more, it would prompt them to address depression or anxiety and perhaps schedule a separate follow-up visit for these concerns, as needed. Although the expert clinician panel's assessments in this study were made by reviewing recorded data measures rather than in-person clinical assessments, the strong correlations with both clinician review and validated measures shows promise in the severity score for clinical use.

Findings from the Bland-Altman analysis revealed that the severity score is significantly different from the PHQ-9, which likely reflects the larger number of measures and the use of weighting for more or less significance for depression or anxiety based on comparison with a large set of individuals. Analysis of outliers in the regression analysis showed instances in which the severity score more accurately identified people with or without clinically significant depression or anxiety than the PHQ-9 or GAD-7 scores, respectively. The identification of the severity score as a measure that is different but still correlates with both validated assessments and clinician review suggests that it can be used as a proxy for those validated measures and that it may have greater utility than those measures alone.

Although combined items from the PHQ-9 and GAD-7 comprise only 16 questions, administering these measures during the primary care visit can be challenging because of time constraints. For example, the addition of just one of these 2 measures could take 20%-40% of the available time of an already short primary care visit lasting 17-21 minutes. Rather than shortening the time needed for measurement, mHealth platforms such as NeuroFlow can aid in MBC integration by taking time for measurement out of the clinical visit. Using the data reported by patients through the use of the app and passively collected data, the severity score provides a measure that is inclusive not only of the full PHQ-9 and GAD-7 data but also other meaningful measures. By assessing a larger number of parameters with an artificial intelligence-driven algorithm, the severity score may be able to identify at-risk individuals who may have been missed with only a PHQ-9 or GAD-7 assessment. The reason for this is the severity score number factoring in multiple assessments and variables, including subjective variables such as mood or sleep. For example, if an individual is not sleeping well, this might be evident in the severity score, whereas it may not be in the PHQ-9 or GAD-7 scores. In addition, digital delivery of the severity score in the patient's electronic health record confers advantages such as HIPAA-compliant screening, mitigation of confidentiality concerns, and completed assessments before meeting with the clinician. This may allow clinicians to use additional time for further assessment and suggestions for treatment, as appropriate.

Providing a system-wide, HIPAA-compliant measurement system that requires minimal training of medical practice personnel can be delivered in a digital manner, and integrating measurement-based care (MBC) into primary care practices confers several advantages. From the clinicians' perspective, the severity score is valuable because it provides a single measure ranging from 1 to 5 that is generated from daily platform use by the patient, can be integrated into the electronic health record, and can alert the clinician to the presence of

concerning endorsements of depression and anxiety symptoms, which in turn can prompt further assessment, treatment, and referrals, as needed. Another advantage of the severity score is that it seamlessly incorporates the measurement of behavioral health symptoms into a person's typical day-to-day activities that include the use of their mobile phones. Taking some time (approximately 30 seconds) to track mood or sleep for multiple days of the week is less onerous for the patient than taking 4 to 8 minutes needed for both the PHQ-9 and GAD-7.

Strengths and Limitations

As with all health-related apps, there is appropriate concern regarding whether individuals will use the app and record data consistently. In this regard, the severity score demonstrated 2 advantages. First, the platform has a high user retention rate of 70%, which is 21 times higher than the typical 30-day retention rate of 3.3% for health-related apps [20]. The platform has user retention rates of 32% and 27% at 6 and 12 months, respectively, and receives a Net Promoter Score of 41, which is 14 points higher than the industry average [26]. This is achieved through well-tested patient engagement techniques that leverage an omni-channel communication strategy. The platform sends patient notifications to register and reminders to engage with the assigned content and activities. Notifications can be in the form of email, SMS text messages, and push notifications depending on the user's preferences. Second, the incorporation of additional patient-reported measures (eg, sleep rating, mood rating, and number of times a measure was tracked) in addition

to PHQ-9 and GAD-7 scores means that a reliable severity measure is available even if 1 measure (eg, the PHQ-9 or GAD-7 score) is not present for a particular individual.

As with all studies, this study is also not without its limitations. For example, given the retrospective design, the sample is not always representative of the population as a whole, and the data set may be at risk of recall bias. Future research with a much larger sample size may prove fruitful, in addition to a prospective study design.

Conclusions

The severity score is a screening measure of the symptoms of depression and anxiety and other important variables, such as mood and sleep, generated with an artificial intelligence-powered algorithm developed from the analysis of 3000 mHealth users. A comparison of over 35,000 severity scores with PHQ-9 and GAD-7 scores taken at the same point in time by the same individuals demonstrated a strong correlation between the severity score and scores from both the PHQ-9 and GAD-7. Clinician reviews of the PHQ-9 and GAD-7 scores also correlate strongly with the severity score. Together, these correlations strongly suggest that the severity score can be used as a proxy measure for the presence of depression and anxiety. However, the Bland-Altman analysis shows that the severity score is a significantly different output measure. Prospective feasibility studies to further measure the sensitivity, specificity, and clinical noninferiority of the severity score are warranted.

Acknowledgments

This study was supported by Small Business Innovation Research funding by the National Science Foundation (NSF 1951062).

Conflicts of Interest

AP and WL are compensated by NeuroFlow, Inc. MLP is a board member of NeuroFlow, Inc.

Multimedia Appendix 1

Multiple linear regression validation.

[PDF File (Adobe PDF File), 11 KB - [formative_v5i11e30313_app1.pdf](#)]

References

1. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. *Cogn Behav Pract* 2015 Feb;22(1):49-59 [FREE Full text] [doi: [10.1016/j.cbpra.2014.01.010](#)] [Medline: [27330267](#)]
2. Masic I, Miokovic M, Muhamedagic B. Evidence based medicine - new approaches and challenges. *Acta Inform Med* 2008;16(4):219-225 [FREE Full text] [doi: [10.5455/aim.2008.16.219-225](#)] [Medline: [24109156](#)]
3. Gondek D, Edbrooke-Childs J, Fink E, Deighton J, Wolpert M. Feedback from outcome measures and treatment effectiveness, treatment efficiency, and collaborative practice: a systematic review. *Adm Policy Ment Health* 2016 May;43(3):325-343 [FREE Full text] [doi: [10.1007/s10488-015-0710-5](#)] [Medline: [26744316](#)]
4. Waldrop J, McGuinness T. Measurement-based care in psychiatry. *J Psychosoc Nurs Ment Health Serv* 2017 Nov 01;55(11):30-35. [doi: [10.3928/02793695-20170818-01](#)] [Medline: [28840934](#)]
5. Lewis C, Boyd M, Puspitasari A, Navarro E, Howard J, Kassab H, et al. Implementing measurement-based care in behavioral health: a review. *JAMA Psychiatry* 2019 Mar 01;76(3):324-335 [FREE Full text] [doi: [10.1001/jamapsychiatry.2018.3329](#)] [Medline: [30566197](#)]
6. Fortney J, Unützer J, Wrenn G, Pyne JM, Smith GR, Schoenbaum M, et al. A tipping point for measurement-based care. *Psychiatr Serv* 2017 Feb 01;68(2):179-188. [doi: [10.1176/appi.ps.201500439](#)] [Medline: [27582237](#)]

7. Zivin K, Miller B, Finke B, Bitton A, Payne P, Stowe EC, et al. Behavioral Health and the Comprehensive Primary Care (CPC) initiative: findings from the 2014 CPC behavioral health survey. *BMC Health Serv Res* 2017 Aug 29;17(1):612 [FREE Full text] [doi: [10.1186/s12913-017-2562-z](https://doi.org/10.1186/s12913-017-2562-z)] [Medline: [28851374](https://pubmed.ncbi.nlm.nih.gov/28851374/)]
8. Alonso J, Liu Z, Evans-Lacko S, Sadikova E, Sampson N, Chatterji S, WHO World Mental Health Survey Collaborators. Treatment gap for anxiety disorders is global: results of the World Mental Health Surveys in 21 countries. *Depress Anxiety* 2018 Mar;35(3):195-208 [FREE Full text] [doi: [10.1002/da.22711](https://doi.org/10.1002/da.22711)] [Medline: [29356216](https://pubmed.ncbi.nlm.nih.gov/29356216/)]
9. The doctor is out: continuing disparities in access to mental and physical health care. National Alliance on Mental Illness. 2017. URL: <https://www.nami.org/Support-Education/Publications-Reports/Public-Policy-Reports/The-Doctor-is-Out/DoctorIsOut> [accessed 2020-08-08]
10. Mental health care services by family physicians. American Association of Family Physicians. URL: <https://www.aafp.org/about/policies/all/mental-services.html> [accessed 2020-06-26]
11. Cavanagh K, Shapiro DA, Van Den Berg S, Swain S, Barkham M, Proudfoot J. The effectiveness of computerized cognitive behavioural therapy in routine care. *Br J Clin Psychol* 2006 Nov;45(Pt 4):499-514. [doi: [10.1348/014466505X84782](https://doi.org/10.1348/014466505X84782)] [Medline: [17076960](https://pubmed.ncbi.nlm.nih.gov/17076960/)]
12. Levine DM, Linder JA, Landon BE. Characteristics of Americans with primary care and changes over time, 2002-2015. *JAMA Intern Med* 2020 Mar 01;180(3):463-466 [FREE Full text] [doi: [10.1001/jamainternmed.2019.6282](https://doi.org/10.1001/jamainternmed.2019.6282)] [Medline: [31841583](https://pubmed.ncbi.nlm.nih.gov/31841583/)]
13. Osborn R, Moulds D, Schneider EC, Doty MM, Squires D, Sarnak DO. Primary care physicians in ten countries report challenges caring for patients with complex health needs. *Health Aff (Millwood)* 2015 Dec;34(12):2104-2112. [doi: [10.1377/hlthaff.2015.1018](https://doi.org/10.1377/hlthaff.2015.1018)] [Medline: [26643631](https://pubmed.ncbi.nlm.nih.gov/26643631/)]
14. National Ambulatory Medical Care Survey National Summary Tables. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2016_namcs [accessed 2021-10-20]
15. Mechanic D, McAlpine DD, Rosenthal M. Are patients' office visits with physicians getting shorter? *N Engl J Med* 2001 Jan 18;344(3):198-204. [doi: [10.1056/NEJM200101183440307](https://doi.org/10.1056/NEJM200101183440307)] [Medline: [11172143](https://pubmed.ncbi.nlm.nih.gov/11172143/)]
16. Tai-Seale M, McGuire TG, Zhang W. Time allocation in primary care office visits. *Health Serv Res* 2007 Oct;42(5):1871-1894 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00689.x](https://doi.org/10.1111/j.1475-6773.2006.00689.x)] [Medline: [17850524](https://pubmed.ncbi.nlm.nih.gov/17850524/)]
17. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med* 2016 Dec 06;165(11):753-760. [doi: [10.7326/M16-0961](https://doi.org/10.7326/M16-0961)] [Medline: [27595430](https://pubmed.ncbi.nlm.nih.gov/27595430/)]
18. Anthes E. Mental health: there's an app for that. *Nature* 2016 Apr 07;532(7597):20-23. [doi: [10.1038/532020a](https://doi.org/10.1038/532020a)] [Medline: [27078548](https://pubmed.ncbi.nlm.nih.gov/27078548/)]
19. Baumel A, Muench F, Edan S, Kane J. 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/)]
20. Parks AC, Williams AL, Tugade MM, Hokes KE, Honomichl RD, Zilca RD. Testing a scalable web and smartphone based intervention to improve depression, anxiety, and resilience: a randomized controlled trial. *Int J Wellbeing* 2018 Dec 08;8(2):22-67. [doi: [10.5502/ijw.v8i2.745](https://doi.org/10.5502/ijw.v8i2.745)]
21. Levis B, Benedetti A, Thombs BD, DEPRESSion Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ* 2019 Apr 09;365:l1476 [FREE Full text] [doi: [10.1136/bmj.l1476](https://doi.org/10.1136/bmj.l1476)] [Medline: [30967483](https://pubmed.ncbi.nlm.nih.gov/30967483/)]
22. Kroenke K, Spitzer R, Williams J. 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/)]
23. Spitzer R, Kroenke K, Williams J, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
24. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)* 2015 Jun 5;25(2):141-151 [FREE Full text] [doi: [10.11613/BM.2015.015](https://doi.org/10.11613/BM.2015.015)] [Medline: [26110027](https://pubmed.ncbi.nlm.nih.gov/26110027/)]
25. Greenberg PE, Fournier A, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry* 2015 Feb;76(2):155-162 [FREE Full text] [doi: [10.4088/JCP.14m09298](https://doi.org/10.4088/JCP.14m09298)] [Medline: [25742202](https://pubmed.ncbi.nlm.nih.gov/25742202/)]
26. What is a good net PROMOTER Score? (2021 NPS Benchmark). Retently. URL: <https://www.retently.com/blog/good-net-promoter-score/> [accessed 2021-10-20]

Abbreviations

- GAD-7:** Generalized Anxiety Disorder-7
- HIPAA:** Health Insurance Portability and Accountability Act
- MBC:** measurement-based care
- mHealth:** mobile health
- PHQ-9:** Personal Health Questionnaire-9

Edited by G Eysenbach; submitted 10.05.21; peer-reviewed by K Mathiasen, T Goldstein; comments to author 17.08.21; revised version received 28.09.21; accepted 02.10.21; published 10.11.21.

Please cite as:

Lynch W, Platt ML, Pardes A

Development of a Severity Score and Comparison With Validated Measures for Depression and Anxiety: Validation Study

JMIR Form Res 2021;5(11):e30313

URL: <https://formative.jmir.org/2021/11/e30313>

doi: [10.2196/30313](https://doi.org/10.2196/30313)

PMID: [34757319](https://pubmed.ncbi.nlm.nih.gov/34757319/)

©William Lynch, Michael L Platt, Adam Pardes. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.11.2021. 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.

Early Report

A Program to Improve Digital Access and Literacy Among Community Stakeholders: Cohort Study

Brittany F Drazich¹, MSN, PhD, RNC; Yeukai Nyikadzino¹, MSN, RN; Kelly T Gleason¹, PhD, RNC

School of Nursing, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:

Brittany F Drazich, MSN, PhD, RNC

School of Nursing

Johns Hopkins University

525 N Wolfe St

Baltimore, MD, 21205

United States

Phone: 1 3023530657

Email: bdrazich@umaryland.edu

Abstract

Background: For many research teams, the role of community stakeholders is critical. However, community stakeholders, especially those in low-income settings, are at risk of being excluded from research and community engagement initiatives during and after the COVID-19 pandemic because of the rapid transition to digital operations.

Objective: We aimed to describe the implementation and feasibility of a program called Addressing the Digital Divide to Improve Patient-Centered Outcomes Research, which was designed to address barriers to technology use, and to examine changes in participants' perceived comfort with digital technology before and after the program.

Methods: To promote full engagement, we worked with 20 existing community leaders to cocreate a training course on using digital technology. We assessed the frequency of technology use and comfort with technology through an adapted 8-item version of the Functional Assessment of Comfort Employing Technology Scale and used the Wilcoxon signed-rank test for survey analysis. We also conducted a focus group session with 10 participants and then performed reflective journaling and content analysis to determine emergent themes.

Results: We found that the program was feasible to implement and worthwhile for participants (15/16, 94%). After the program, the participants perceived an increase in the frequency of technology use ($z=2.76$, $P=.006$). The participants reported that the program was successful because of the technology training program, but recommended that the program have a slower pace and include a helpline number that they could call with questions.

Conclusions: Future programs should consider that populations with low literacy view technology training as a core element to decreasing technology disparity. This study demonstrates that through low-cost input, community members can be provided the resources and training needed to virtually participate in research studies or community engagement initiatives.

(*JMIR Form Res* 2021;5(11):e30605) doi:[10.2196/30605](https://doi.org/10.2196/30605)

KEYWORDS

technology; disparities; digital access; digital literacy; community; stakeholders; digital health; digital divide; patient-centered outcomes

Introduction

In an effort to mitigate the spread of disease and prevent unnecessary death caused by the COVID-19 pandemic, public health leaders and policymakers globally have recommended or instituted physical distancing measures [1-4]. Subsequently, people have replaced in-person interactions with virtual interactions [5-9]. For example, in the United States, the use of

telehealth, virtual learning, telework, and personal video chat communication has increased [10-13]. This use of technology has allowed people to remain active and engaged with others during the pandemic.

Research teams, many of which include community members, have also adapted operations to virtual platforms such as videoconferencing wherever possible. For many research teams, the role of community stakeholders is critical. Community

stakeholders can assess the importance and cultural relevance of a research question. Involvement of community stakeholders in research can improve recruitment and sustainability of outcomes [14]. From an ethical standpoint, many assert that the community has a right to be aware of and involved in research that could affect them [15]. Community stakeholders involved in research are often leaders in their communities who facilitate connections and ensure awareness of community resources [15-18].

Although technology has the potential to assist community stakeholders in remaining engaged during the COVID-19 pandemic, there are technology-based disparities [19-21]. Many community stakeholders, especially those in low-income settings, are at risk of being excluded from research and community engagement initiatives during and after the COVID-19 pandemic because of the rapid transition to digital operations. In Baltimore, Maryland where this study was conducted, an estimated 40% of households do not have wireless internet service and one-third of households do not have a desktop or laptop computer [22-24]. Black Americans, who make up 67% of the population of Baltimore [25], experience more barriers to technology use than the rest of the city's population because of systemic and structural racism that has led to socioeconomic disparities [26,27]. Black older adults experience additional barriers to technology use such as lack of guidance, lack of confidence, limited resources, and the perception that technology is complex [28,29].

These barriers to technology use severely restrict Baltimore's community members from getting involved in research teams and supporting their communities during the COVID-19 pandemic. It is essential to address technology-based disparities, or the "digital divide," to ensure full community engagement for patient-centered research design and generalizable research samples both during and after this pandemic. This study aimed to describe a program called Addressing the Digital Divide to Improve Patient-Centered Outcomes Research or ADD2PCOR, which was created to address the barriers to technology use. Through this program, we aimed to provide a time-efficient, cost-effective, and feasible training course to bridge the digital divide among Baltimore's community stakeholders and impart technological knowledge and tools to those who would otherwise be unable or unlikely to participate digitally. In this paper, we describe the implementation and feasibility of ADD2PCOR and examine participants' comfort with digital technology before and after the program.

Methods

Project Description

ADD2PCOR Goals

We co-designed, implemented, and refined ADD2PCOR to provide community stakeholders and vulnerable patients with basic technological knowledge and equipment that would help them to participate virtually. We intended for this program to be both cost- and time-efficient so that it could be broadly implemented by research teams, both within our institution and

elsewhere, who wish to engage people with low technology literacy in research activities.

Participants and Recruitment

We recruited participants who were 18 years or older, participated as a community member on a research advisory council or similar committee or confirmed intent to join a research advisory council or similar committee in the coming year, and verbalized that they did not have regular access to internet at their home or a digital tool beyond their phone to access the internet. We performed snowball sampling through our existing partnerships with patients and the community and recruited 20 community stakeholders.

Training Course Development

To promote full engagement, we worked with existing community leaders to cocreate a training course on using digital technology. Brevity and feasibility were emphasized throughout training course creation to ensure that other organizations may be able to easily implement the course without the need to devote significant resources. We obtained input from community and patient groups such as the Community Research Advisory Council, the Patient and Family Advisory Council, and Patients Aligned with Research Teams and ER Nurses to Improve Diagnosis to understand the difficulties that patients and community members have faced in engaging in virtual meetings and using other basic technologies. Training course development in this project followed the classic 4 elements of the Tyler Model of planning, designing, implementing, and evaluating with community stakeholders engaged in the conception of each aspect [30]. Codevelopment is the cornerstone of the design portion of this curriculum development. We created written material to loosely guide training course progression for both the trainer and community stakeholders. All learning topics in the training course were determined based on community stakeholder input, which were then amended based on individual participant needs during program implementation.

Project Implementation

Each community stakeholder recruited for this study received an Amazon Fire tablet, which was purchased at a price as low as \$40. If they did not already have broadband internet access in their home, they were given a 1-year Comcast Internet Essentials plan, which was available for \$9.95 a month inclusive of setup and rental costs [31]. The maximum cost of providing both internet access and a tablet was \$200 a year, which is comparable to that spent on meals and parking for participants in many research or engagement programs.

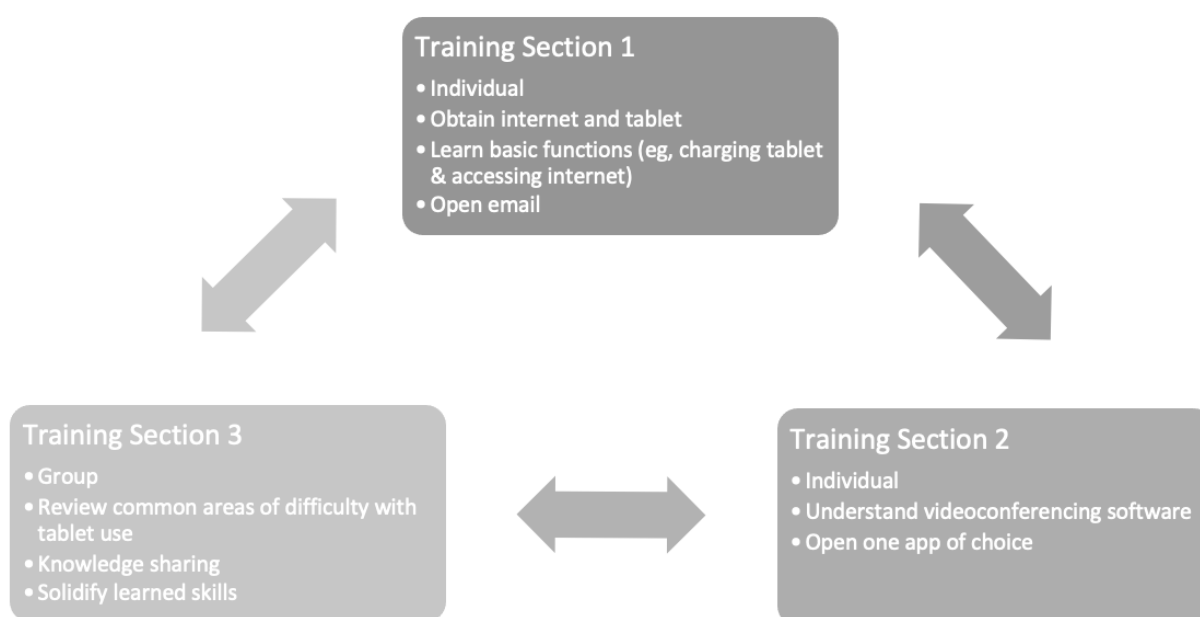
Based on suggestions from community stakeholders, the first section of the training course focused on ensuring that participants were comfortable with the basic functions of the tablet, including switching on the tablet, charging the tablet, and connecting to the internet. Once we confirmed that participants were able to access the internet, we introduced free basic services such as Gmail. We ensured that participants could open their email accounts as well as documents and calendar invitations sent through email. The second section of the training course focused on ensuring that participants were comfortable using videoconferencing and at least one application (or "app")

of their choosing. For example, many participants requested assistance in learning how to access their patient portal for health information or use social media to view pictures of their grandchildren. After the majority of participants were confident with both connecting to the internet and videoconferencing, we conducted 1-hour video group meetings to review common areas of difficulty with the use of the tablet. Participants who had not yet mastered videoconferencing joined the video meetings via phone-based audio. These group meetings were opportunities for the community stakeholder participants to learn from each other or solidify learned skills.

We compensated participants \$30 for each hour of training that they attended. Although the training course was originally planned as in-person instruction, we delivered all training sessions virtually because of pandemic-related safety concerns.

Similarly, although the training course was originally planned as group-based learning, the majority of the training course was delivered through one-on-one instruction over phone or video chat. We determined one-on-one instruction to be more efficient following the switch to virtual instruction because it allowed the delivery of individualized education and limited extraneous background noise during video meetings. Although we originally planned for the training course to be 3 hours long for all participants, we changed this to a more flexible format during the program based on participant needs and feedback. However, on average, participants completed 3 hours of training course content. Although the training course components and pace were individualized based on the technology literacy and preferences of each community stakeholder, a step-by-step example is described in [Figure 1](#).

Figure 1. Example of a community stakeholder training course. The arrows indicate the iterative nature of the training course.



Data Collection and Analysis

We used Qualtrics, an online survey platform, to understand the frequency of use of various digital technologies by community stakeholders before and after ADD2PCOR. We considered frequency of technology use to be an indicator of comfort with technology. We assessed frequency of technology use through an adapted 8-item version of the Functional Assessment of Comfort Employing Technology Scale (FACETS). FACETS encompasses 5 functional domains: social, e-commerce, technical, health care, and home [32]. An example survey question is “I use Google or another search engine to find answers to questions.” Scores for participants’ responses to each question ranged from 0 to 6, with higher scores indicating a greater frequency of technology use [32]. The survey responses were combined for an overall score ranging from 0 to 48. The adapted FACETS survey is shown in [Multimedia Appendix 1](#). We used the Wilcoxon signed-rank test to compare responses on the frequency of technology use before and after ADD2PCOR, considering a *P* value <.05 as significant.

After all community stakeholders participated in the training course, we conducted a focus group session with the ADD2PCOR participants who expressed an interest in joining it. The purpose of the focus group was to receive feedback from the participants on their experience with ADD2PCOR. The researcher who moderated the focus group discussion was an experienced qualitative interviewer. In an effort to prevent bias, the moderator and notetaker present had no prior interactions with the ADD2PCOR participants. Questions in the semistructured interview guide were related to previous comfort with and use of digital technology, motivation to join ADD2PCOR, positive and negative feedback on ADD2PCOR, opinion on the training aspects of the program, and comfort with and use of digital technology after ADD2PCOR. The focus group session was conducted as a virtual video meeting of 10 participants that lasted approximately 1 hour and 15 minutes.

Although the focus group meeting was not recorded, thorough notes of direct quotations were taken from participants throughout the session, followed by reflective journaling and content analysis conducted in a manner similar to that by

Halcomb and Davidson [33]. We selected content analysis for the organization of focus group data because of our study's focus on obtaining objective feedback for program improvement. The focus group moderator (BFD) utilized the coding software f4analyze (audiotranskription) to organize direct quotations, and the project lead (KTG) reviewed the content analysis and field notes and validated the selected themes [34].

Results

Characteristics of the Sample

In total, 16 of the 20 community stakeholder participants (85%) completed the demographical survey (Table 1) and the FACETS survey (Table 2). The majority of the participants were female and aged 65-74 years. All study participants identified as Black individuals. The majority of participants were extremely satisfied with the program (11/16, 69%) and strongly agreed that the experience was worthwhile (15/16, 94%).

Table 1. Sample characteristics.

Characteristic	Participants (N=17), n (%)
Age (years)	
45-54	1 (6)
55-64	3 (18)
65-74	10 (59)
75-84	2 (12)
>85	1 (6)
Race	
Black	17 (100)
Gender	
Women	14 (82)
Men	3 (18)
Education level	
Less than high school	2 (13)
High school graduate	4 (25)
Some college	6 (38)
4-year degree or more	5 (29)
Satisfaction level^a	
Extremely satisfied	11 (69)
Somewhat satisfied	5 (31)
How likely recommend to a friend^a	
Extremely likely	14 (88)
Somewhat likely	2 (12)
Found experience worthwhile^a	
Strongly agree	15 (94)
Somewhat agree	1 (7)

^an=16.

Table 2. Changes in technology use frequency (N=17).

FACETS ^a domains	Score, mean (SD)	z value	P value
Texting frequency		-2.18	.03
Pre	4.15 (1.90)		
Post	4.88 (1.68)		
Social media use frequency		-1.23	.22
Pre	2.25 (1.62)		
Post	3.38 (1.89)		
Wi-Fi use frequency		-2.75	.01
Pre	3.25 (1.62)		
Post	5.13 (1.41)		
Videoconferencing frequency		-2.73	.01
Pre	2.45 (1.73)		
Post	4.81 (1.64)		
Searching on Google frequency		-2.17	.03
Pre	4.35 (1.63)		
Post	5 (1.80)		
Opening files frequency		-1.04	.30
Pre	2.45 (1.76)		
Post	3.62 (2.30)		
Opening shared files frequency		-2.04	.04
Pre	1.90 (1.86)		
Post	3.5 (2.34)		
Opening calendar invites frequency		-3.20	.001
Pre	1.45 (1.39)		
Post	4.19 (2.04)		
Combined frequency score		-2.76	.006
Pre	24.40 (11.05)		
Post	38.31 (10.64)		

^aFACETS: Functional Assessment of Comfort Employing Technology Scale.

Survey Results for Technology Use

Cronbach α was .88 for the adapted FACETS survey. After comparing the FACETS measure before and after ADD2PCOR, we found a significant increase in the score for the following 6 subdomains of frequency of technology use: texting, Wi-Fi, videoconferencing, opening files sent by others, searching on Google, and opening calendar invites. The subdomains with the greatest increase in score on a scale of 0-10 were “videoconferencing frequency,” with an average increase of 2.36, and “opening calendar invites frequency,” with an average increase of 2.74. A comparison of the combined total subdomains scores before and after the program showed an increase in score by 13.91 (scale 0-48) for the frequency of technology use.

Focus Group Results

During focus group analysis, we identified 3 major themes: motivation and benefits of the program, training as the core to success, and areas for program improvement. The findings are summarized below.

Motivation and Benefits of the Program

The participants overwhelmingly joined ADD2PCOR because they wanted “to learn.” Some participants reported that they generally “like to learn new things,” whereas others perceived a specific gap in knowledge regarding technology. The participants discussed that having skills in technology gives them comfort because technology can provide entertainment or enable them to contact others when in need. For example, one participant stated,

[Without this program], I would be where I was before I started (laughs). You have broadened my

resource base, gave me more tools to work with, to navigate the pandemic with.

Some of the participants reported that they had limited knowledge on affordable internet and device options before the program and that they intended to continue with the internet plan through the year after study conclusion. The participants were especially appreciative of being able to video chat with loved ones who they would otherwise not be able to see face-to-face because of the pandemic or geographical distance. Finally, the participants said that the program allows them to better serve their community during the pandemic. For example, one participant stated that,

[The program] allows us to dig into more resources for the community. People are isolated during the pandemic... We are community people. We want to continue doing what we do – getting a hold of resources and sharing them. The ultimate goal is helping others.

Training as the Core to Success

Most participants reported that the most valuable aspect of the program was the training or classes. The participants described how ADD2PCOR would be “useless” without the training classes, as their tablet “would just be lying there.” One participant narrated their experience of acquiring a new iPad, only to later give it away because he did not understand how to use it. Another participant described how her family members use technology on her behalf instead of offering to teach her how to use it. The participants expressed appreciation for having an instructor for the tablet who is patient with basic questions or slow progress. For example, one participant stated, “y’all take your time and take me step by step to learn the steps without getting frustrated.” The participants stated that receiving the tablet and receiving training to use it “go hand in glove”; each is imperative for the success of the other.

Areas for Program Improvement

The survey results indicated that the participants were extremely satisfied with ADD2PCOR; however, they did make some suggestions for similar future programs during the focus group session. Some participants suggested a slower pace of training “so we can learn every icon that pops up on our program and how to use it, what your tablet can do.” Another participant suggested that periodic recaps on past learned skills be offered for adults aged 80 years and older. Two older participants suggested that older adults, especially those with poor eyesight, be taught to use audio virtual assistants such as Alexa or audiobooks. The participants also expressed the importance of a helpline to call with questions outside of formal training meetings. Similarly, one participant recommended that veteran participants be partnered with newer participants to help troubleshoot problems with basic tablet functions.

Discussion

Addressing disparities in technology access and use among community stakeholders is essential to creating and implementing culturally competent research studies and interventions. ADD2PCOR addressed this need by providing

community stakeholders with a tablet and internet as needed and by implementing a self-paced technology training course. We found that ADD2PCOR was feasible to implement and worthwhile (15/16, 94%) for participants. After the program, the participants perceived an increase in the frequency of technology use in 6 subdomains: texting, Wi-Fi, videoconferencing, opening files sent by others, searching using Google, and opening calendar invites. The participants largely joined the program “to learn,” and believed that the greatest benefit of the program was that the skills they learned will help them obtain and provide help to others. The participants reported that the program was successful because of the technology training program, but recommended that the pace of the program be reduced and that a helpline number that they could call with questions be provided.

Considering the small sample size of this study, it is promising that participants’ frequency of technology use significantly improved in the majority of tested subdomains. The subdomains that showed a significant increase in score were those of core skills taught during the training or practiced by participants throughout the study, except for text messaging. Improving comfort with the use of technology in other areas may have translated to improvement in comfort with text messaging. These skills that showed a significant increase in score are critical skills needed for the community stakeholders to continue their involvement in research projects or community engagement in the virtual environment. While social media use increased among participants, there was no significant increase in the score for this subdomain. During the intervention, the training focused more on basic tablet functions than on the use of social media, per participant request. During the focus group session, the suggestion that the pace of training be reduced was provided mainly by older participants. Future studies on technology training might benefit from creating homogenous subgroups for training divided by baseline technology knowledge. Except for a few reports [35,36], most studies on technology-based interventions have focused on either training or improving access [37-39]. Our study results suggest that both are core elements for improving comfort with and use of technology and should be addressed in tandem.

Most of the participants in this study were aged 65 years or older and identified as Black individuals, which reflects the demographics of populations experiencing disparity in technology use and access in the United States. Many older Black Americans experience intersectional barriers to technology use such as inexperience arising from age-related preferences or disability [20,40] or financial constraints related to structural inequities [41]. During the COVID-19 pandemic, such technology disparities might contribute to poor vaccine uptake or limited health care options for these populations [42]. For example, COVID-19 vaccine scheduling and telehealth appointments are both health care activities that require technology-based literacy. Technology should be used as a tool to decrease and not perpetuate disparities [38,39,43-45].

During the pandemic, studies are being adapted to virtual platforms and this will likely continue after the pandemic. It is essential that all people have the resources and knowledge needed to engage in such virtual studies. Broadband internet

and device ownership is less common in certain populations such as older adults or low-income populations [46,47]. Unfortunately, many virtual studies and interventions, even those aimed at decreasing disparities, require such resources for participation [37,45,48]. This study demonstrates that research teams can provide participants with a tablet, internet, and training through low-cost input. The intention of some of the participants to continue with the internet plan through the year after study conclusion indicates the sustainability of the program.

The major limitations of this study were the sample size, sampling procedure, and groupthink. The small sample size was appropriate for this feasibility study but might have had limited power in detecting smaller differences. In addition, we used a snowball sampling technique that might have included participants that are not representative of other community stakeholders in Baltimore. Finally, even though the focus group session was moderated by a researcher who had no prior interaction with the participants, participant feedback was overwhelmingly positive. The participants might have felt pressured to respond in a manner similar to their peers (groupthink) [49]. Regardless of these limitations, this study has many strengths. With the inclusion of community leaders and existing collaborations, we were able to recruit people who might have otherwise been difficult to reach through research. Furthermore, this study was able to implement participation feedback at all stages, thus providing individualized training as needed and making adjustments when components were found to be unsuccessful.

The findings of this study highlight areas in need of future research and policy change. This study demonstrates a feasible intervention that improves comfort with and use of technology for community stakeholders. Many previous researchers have excluded otherwise eligible participants from studies because of a lack of internet access or devices needed to participate in the study. The results of this study indicate that through low-cost input, community members can be provided such resources and be included in technology-based studies. Future studies could implement a program similar to ADD2PCOR in a larger sample and integrate the constructive feedback presented in the focus group. This study suggests that providing the necessary technological equipment alone might not be sufficient for improving technology use in all populations; researchers must also consider a participant's sociotechnical environment [50,51]. For example, this study's sample of older Black Americans with low technology literacy viewed technology training and a technology helpline as important supports for technology use. Similarly, if researchers or organizations intend to provide group learning for community members or study participants, they should consider dividing groups by technology literacy so that the training pace complements each individual. Recently, the Federal Communications Commission began the Emergency Broadband Benefit program, which provides affordable broadband services to low-income Americans [52]. Researchers who work with community stakeholders could also explore currently available free online classes that are specifically designed to improve technology literacy. Including community members as leaders or participants in studies is critical, and the lack of internet, technological equipment, or technology literacy are modifiable factors that can be addressed by research teams.

Acknowledgments

KTG receives funding from the following sources: The National Institutes of Health's (NIH) National Center for Advancing Translational Sciences (NCATS) Institutional Career Development Core (grant number KL2 TR003099), NIH NCATS Johns Hopkins Institute for Clinical and Translational Research (grant number UL1TR003098), Patient-Centered Outcomes Research Institute (award number EAIN-00178), Addressing the Digital Divide to Improve Patient Centered-Outcomes Research, and the Gordon and Betty Moore Foundation, #9904, Development of a Patient-Reported Measure Set of Diagnostic Excellence. BFD receives funding from the Robert Wood Johnson Foundation Future of Nursing Scholars Program.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Adapted Functional Assessment of Comfort Employing Technology Scale survey.

[[DOCX File, 14 KB - formative_v5i11e30605_app1.docx](#)]

References

1. Adhikari SP, Meng S, Wu Y, Mao Y, Ye R, Wang Q, et al. Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infect Dis Poverty* 2020 Mar 17;9(1):29 [[FREE Full text](#)] [doi: [10.1186/s40249-020-00646-x](https://doi.org/10.1186/s40249-020-00646-x)] [Medline: [32183901](https://pubmed.ncbi.nlm.nih.gov/32183901/)]
2. Nicola M, O'Neill N, Sohrabi C, Khan M, Agha M, Agha R. Evidence based management guideline for the COVID-19 pandemic - Review article. *Int J Surg* 2020 May;77:206-216 [[FREE Full text](#)] [doi: [10.1016/j.ijssu.2020.04.001](https://doi.org/10.1016/j.ijssu.2020.04.001)] [Medline: [32289472](https://pubmed.ncbi.nlm.nih.gov/32289472/)]
3. COVID-19 - social distancing. Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html> [accessed 2021-05-15]

4. Thu TPB, Ngoc PNH, Hai NM, Tuan LA. Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. *Sci Total Environ* 2020 Nov 10;742:140430 [FREE Full text] [doi: [10.1016/j.scitotenv.2020.140430](https://doi.org/10.1016/j.scitotenv.2020.140430)] [Medline: [32623158](https://pubmed.ncbi.nlm.nih.gov/32623158/)]
5. Brynjolfsson E, Horton J, Ozimek A, Rock D, Sharma G, TuYe H. COVID-19 and remote work: An early look at US data. National Bureau of Economic Research. 2020. URL: <https://www.nber.org/papers/w27344> [accessed 2021-07-15]
6. Adnan M. Online learning amid the COVID-19 pandemic: Students perspectives. *JSPS* 2020 Jun 22;1(2):45-51. [doi: [10.33902/jpsp.2020261309](https://doi.org/10.33902/jpsp.2020261309)]
7. Basilaia G, Kvavadze D. Transition to Online Education in Schools during a SARS-CoV-2 Coronavirus (COVID-19) Pandemic in Georgia. *Pedagogical Res* 2020;5(4):em0060. [doi: [10.29333/pr/7937](https://doi.org/10.29333/pr/7937)]
8. Week 16 Household Pulse Survey: September 30 – October 12. United States Census Bureau. 2020. URL: <https://www.census.gov/data/tables/2020/demo/hhp/hhp16.html> [accessed 2021-04-28]
9. Belzunegui-Eraso A, Erro-Garcés A. Teleworking in the Context of the Covid-19 Crisis. *Sustainability* 2020 May 01;12(9):3662. [doi: [10.3390/su12093662](https://doi.org/10.3390/su12093662)]
10. Mann D, Chen J, Chunara R, Testa P, Nov O. COVID-19 transforms health care through telemedicine: Evidence from the field. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: [10.1093/jamia/ocaa072](https://doi.org/10.1093/jamia/ocaa072)] [Medline: [32324855](https://pubmed.ncbi.nlm.nih.gov/32324855/)]
11. Mukhopadhyay S, Booth A, Calkins S, Doxtader EE, Fine SW, Gardner JM, et al. Leveraging Technology for Remote Learning in the Era of COVID-19 and Social Distancing. *Arch Pathol Lab Med* 2020 Sep 01;144(9):1027-1036 [FREE Full text] [doi: [10.5858/arpa.2020-0201-ED](https://doi.org/10.5858/arpa.2020-0201-ED)] [Medline: [32364793](https://pubmed.ncbi.nlm.nih.gov/32364793/)]
12. Prigoff J, Hunter M, Nowygrod R. Medical Student Assessment in the Time of COVID-19. *J Surg Educ* 2021 Mar;78(2):370-374 [FREE Full text] [doi: [10.1016/j.jsurg.2020.07.040](https://doi.org/10.1016/j.jsurg.2020.07.040)] [Medline: [32819868](https://pubmed.ncbi.nlm.nih.gov/32819868/)]
13. Lee M, Zhao J, Sun Q, Pan Y, Zhou W, Xiong C, et al. Human mobility trends during the early stage of the COVID-19 pandemic in the United States. *PLoS One* 2020 Nov 9;15(11):e0241468 [FREE Full text] [doi: [10.1371/journal.pone.0241468](https://doi.org/10.1371/journal.pone.0241468)] [Medline: [33166301](https://pubmed.ncbi.nlm.nih.gov/33166301/)]
14. Jagosh J, Macaulay A, Pluye P, Salsberg J, Bush PL, Henderson J, et al. Uncovering the benefits of participatory research: implications of a realist review for health research and practice. *Milbank Q* 2012 Jun;90(2):311-346 [FREE Full text] [doi: [10.1111/j.1468-0009.2012.00665.x](https://doi.org/10.1111/j.1468-0009.2012.00665.x)] [Medline: [22709390](https://pubmed.ncbi.nlm.nih.gov/22709390/)]
15. Frisby W, Reid C, Millar S, Hoerber L. Putting “participatory” into participatory forms of action research. *J Sport Manag* 2005;19(4):367-386. [doi: [10.1123/jism.19.4.367](https://doi.org/10.1123/jism.19.4.367)]
16. Michener L, Cook J, Ahmed SM, Yonas MA, Coyne-Beasley T, Aguilar-Gaxiola S. Aligning the Goals of Community-Engaged Research. *Academic Medicine* 2012;87(3):285-291. [doi: [10.1097/acm.0b013e3182441680](https://doi.org/10.1097/acm.0b013e3182441680)]
17. Joosten YA, Israel TL, Williams NA, Boone LR, Schlundt DG, Mouton CP, et al. Community Engagement Studios. *Academic Medicine* 2015;90(12):1646-1650. [doi: [10.1097/acm.0000000000000794](https://doi.org/10.1097/acm.0000000000000794)]
18. Cunningham-Erves J, Mayo-Gamble T, Vaughn Y, Hawk J, Helms M, Barajas C, et al. Engagement of community stakeholders to develop a framework to guide research dissemination to communities. *Health Expect* 2020 Aug 25;23(4):958-968 [FREE Full text] [doi: [10.1111/hex.13076](https://doi.org/10.1111/hex.13076)] [Medline: [32452067](https://pubmed.ncbi.nlm.nih.gov/32452067/)]
19. Ramsetty A, Adams C. Impact of the digital divide in the age of COVID-19. *J Am Med Inform Assoc* 2020 Jul 01;27(7):1147-1148 [FREE Full text] [doi: [10.1093/jamia/ocaa078](https://doi.org/10.1093/jamia/ocaa078)] [Medline: [32343813](https://pubmed.ncbi.nlm.nih.gov/32343813/)]
20. Friemel TN. The digital divide has grown old: Determinants of a digital divide among seniors. *New Media & Society* 2014 Jun 12;18(2):313-331. [doi: [10.1177/1461444814538648](https://doi.org/10.1177/1461444814538648)]
21. Lai J, Widmar NO. Revisiting the Digital Divide in the COVID-19 Era. *Appl Econ Perspect Policy* 2020 Oct 12;43(1):458-464 [FREE Full text] [doi: [10.1002/aapp.13104](https://doi.org/10.1002/aapp.13104)] [Medline: [33230409](https://pubmed.ncbi.nlm.nih.gov/33230409/)]
22. Horrigan B. Baltimore's Digital Divide: Gaps in Internet Connectivity and the Impact on Low-income City Residents. Abell Foundation. 2020. URL: <https://abell.org/publications/baltimores-digital-divide-gaps-internet-connectivity-and-impact-low-income-city> [accessed 2021-04-11]
23. Anderson M, Kumar M. Digital divide persists even as lower-income Americans make gains in tech adoption. Pew Research Center. 2019. URL: <https://www.pewresearch.org/fact-tank/2019/05/07/digital-divide-persists-even-as-lower-income-americans-make-gains-in-tech-adoption/> [accessed 2021-05-10]
24. Jensen JD, King AJ, Davis LA, Guntzwiller LM. Utilization of internet technology by low-income adults: the role of health literacy, health numeracy, and computer assistance. *J Aging Health* 2010 Sep 21;22(6):804-826. [doi: [10.1177/0898264310366161](https://doi.org/10.1177/0898264310366161)] [Medline: [20495159](https://pubmed.ncbi.nlm.nih.gov/20495159/)]
25. Quickfacts: Baltimore City (county), Maryland. United States Census Bureau. 2020. URL: <https://www.census.gov/quickfacts/fact/table/baltimorecitycountymaryland/RHI225219#qf-headnote-a> [accessed 2021-04-11]
26. Price Z. Race-Based Health Disparities and the Digital Divide: Implications for Nursing Practice. *J Natl Black Nurses Assoc* 2015 Dec;26(2):79-83. [Medline: [27045161](https://pubmed.ncbi.nlm.nih.gov/27045161/)]
27. 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/)]

28. Vaportzis E, Clausen MG, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. *Front Psychol* 2017 Oct 04;8:1687 [FREE Full text] [doi: [10.3389/fpsyg.2017.01687](https://doi.org/10.3389/fpsyg.2017.01687)] [Medline: [29071004](https://pubmed.ncbi.nlm.nih.gov/29071004/)]
29. Di Giacomo D, Ranieri J, D'Amico M, Guerra F, Passafiume D. Psychological Barriers to Digital Living in Older Adults: Computer Anxiety as Predictive Mechanism for Technophobia. *Behav Sci (Basel)* 2019 Sep 11;9(9):96 [FREE Full text] [doi: [10.3390/bs9090096](https://doi.org/10.3390/bs9090096)] [Medline: [31514364](https://pubmed.ncbi.nlm.nih.gov/31514364/)]
30. Tyler R. Basic principles of curriculum and instruction. Chicago, IL: University of Chicago Press; 2013:1-144.
31. Affordable internet at home for eligible households. Internet Essentials from COMCAST. 2021. URL: <https://www.internetessentials.com> [accessed 2021-05-18]
32. Lepkowsky C, Arndt S. Functional assessment of currently employed technology scale (FACETS): Reliability and validity. *Int J Med Sci Clin Invent* 2018;5(9):4064-4068. [doi: [10.18535/ijmsci/v5i9.07](https://doi.org/10.18535/ijmsci/v5i9.07)]
33. Halcomb EJ, Davidson PM. Is verbatim transcription of interview data always necessary? *Appl Nurs Res* 2006 Feb;19(1):38-42. [doi: [10.1016/j.apnr.2005.06.001](https://doi.org/10.1016/j.apnr.2005.06.001)] [Medline: [16455440](https://pubmed.ncbi.nlm.nih.gov/16455440/)]
34. f4analyse. audiotranskription. 2020. URL: <https://www.audiotranskription.de/english/f4-analyse> [accessed 2021-05-18]
35. Delello JA, McWhorter RR. Reducing the Digital Divide: Connecting Older Adults to iPad Technology. *J Appl Gerontol* 2017 Jan 09;36(1):3-28. [doi: [10.1177/0733464815589985](https://doi.org/10.1177/0733464815589985)] [Medline: [26084479](https://pubmed.ncbi.nlm.nih.gov/26084479/)]
36. Gilson A, Dodds D, Kaur A, Potteiger M, Ford Ii JH. Using Computer Tablets to Improve Moods for Older Adults With Dementia and Interactions With Their Caregivers: Pilot Intervention Study. *JMIR Form Res* 2019 Sep 03;3(3):e14530 [FREE Full text] [doi: [10.2196/14530](https://doi.org/10.2196/14530)] [Medline: [31482847](https://pubmed.ncbi.nlm.nih.gov/31482847/)]
37. Toscos T, Drouin M, Pater J, Flanagan M, Pfafman R, Mirro M. Selection biases in technology-based intervention research: patients' technology use relates to both demographic and health-related inequities. *J Am Med Inform Assoc* 2019 Aug 01;26(8-9):835-839 [FREE Full text] [doi: [10.1093/jamia/ocz058](https://doi.org/10.1093/jamia/ocz058)] [Medline: [31172193](https://pubmed.ncbi.nlm.nih.gov/31172193/)]
38. Lyles C, Schillinger D, Sarkar U. Connecting the Dots: Health Information Technology Expansion and Health Disparities. *PLoS Med* 2015 Jul 14;12(7):e1001852 [FREE Full text] [doi: [10.1371/journal.pmed.1001852](https://doi.org/10.1371/journal.pmed.1001852)] [Medline: [26172977](https://pubmed.ncbi.nlm.nih.gov/26172977/)]
39. Satariano WA, Scharlach AE, Lindeman D. Aging, place, and technology: toward improving access and wellness in older populations. *J Aging Health* 2014 Dec 11;26(8):1373-1389. [doi: [10.1177/0898264314543470](https://doi.org/10.1177/0898264314543470)] [Medline: [25502245](https://pubmed.ncbi.nlm.nih.gov/25502245/)]
40. Gell NM, Rosenberg DE, Demiris G, LaCroix AZ, Patel KV. Patterns of technology use among older adults with and without disabilities. *Gerontologist* 2015 Jun;55(3):412-421 [FREE Full text] [doi: [10.1093/geront/gnt166](https://doi.org/10.1093/geront/gnt166)] [Medline: [24379019](https://pubmed.ncbi.nlm.nih.gov/24379019/)]
41. Gordon NP, Banegas MP, Tucker-Seeley RD. Racial-ethnic differences in prevalence of social determinants of health and social risks among middle-aged and older adults in a Northern California health plan. *PLoS One* 2020 Nov 4;15(11):e0240822 [FREE Full text] [doi: [10.1371/journal.pone.0240822](https://doi.org/10.1371/journal.pone.0240822)] [Medline: [33147232](https://pubmed.ncbi.nlm.nih.gov/33147232/)]
42. Painter EM, Ussery EN, Patel A, Hughes MM, Zell ER, Moulia DL, et al. Demographic Characteristics of Persons Vaccinated During the First Month of the COVID-19 Vaccination Program - United States, December 14, 2020-January 14, 2021. *MMWR Morb Mortal Wkly Rep* 2021 Feb 05;70(5):174-177 [FREE Full text] [doi: [10.15585/mmwr.mm7005e1](https://doi.org/10.15585/mmwr.mm7005e1)] [Medline: [33539333](https://pubmed.ncbi.nlm.nih.gov/33539333/)]
43. Rivers BM, Bernhardt JM, Fleisher L, Green BL. Opportunities and challenges of using technology to address health disparities. *Future Oncol* 2014 Mar;10(4):519-522. [doi: [10.2217/fon.14.17](https://doi.org/10.2217/fon.14.17)] [Medline: [24754582](https://pubmed.ncbi.nlm.nih.gov/24754582/)]
44. Ye S, Kronish I, Fleck E, Fleischut P, Homma S, Masini D, et al. Telemedicine Expansion During the COVID-19 Pandemic and the Potential for Technology-Driven Disparities. *J Gen Intern Med* 2021 Jan 26;36(1):256-258 [FREE Full text] [doi: [10.1007/s11606-020-06322-y](https://doi.org/10.1007/s11606-020-06322-y)] [Medline: [33105000](https://pubmed.ncbi.nlm.nih.gov/33105000/)]
45. Zhang X, Hailu B, Tabor D, Gold R, Sayre MH, Sim I, et al. Role of Health Information Technology in Addressing Health Disparities: Patient, Clinician, and System Perspectives. *Med Care* 2019 Jun;57 Suppl 6 Suppl 2:S115-S120 [FREE Full text] [doi: [10.1097/MLR.0000000000001092](https://doi.org/10.1097/MLR.0000000000001092)] [Medline: [31095049](https://pubmed.ncbi.nlm.nih.gov/31095049/)]
46. Huxhold O, Hees E, Webster NJ. Towards bridging the grey digital divide: changes in internet access and its predictors from 2002 to 2014 in Germany. *Eur J Ageing* 2020 Sep 10;17(3):271-280 [FREE Full text] [doi: [10.1007/s10433-020-00552-z](https://doi.org/10.1007/s10433-020-00552-z)] [Medline: [32904732](https://pubmed.ncbi.nlm.nih.gov/32904732/)]
47. Levine DM, Lipsitz SR, Linder JA. Trends in Seniors' Use of Digital Health Technology in the United States, 2011-2014. *JAMA* 2016 Aug 02;316(5):538-540. [doi: [10.1001/jama.2016.9124](https://doi.org/10.1001/jama.2016.9124)] [Medline: [27483069](https://pubmed.ncbi.nlm.nih.gov/27483069/)]
48. van Veen T, Binz S, Muminovic M, Chaudhry K, Rose K, Calo S, et al. Potential of Mobile Health Technology to Reduce Health Disparities in Underserved Communities. *West J Emerg Med* 2019 Aug 06;20(5):799-802 [FREE Full text] [doi: [10.5811/westjem.2019.6.41911](https://doi.org/10.5811/westjem.2019.6.41911)] [Medline: [31539337](https://pubmed.ncbi.nlm.nih.gov/31539337/)]
49. McLeod K, Feller E. Groupthink: What's So Great About Teamwork? *R I Med J (2013)* 2019 Sep 03;102(7):8-9. [Medline: [31480809](https://pubmed.ncbi.nlm.nih.gov/31480809/)]
50. Timmermans S, Berg M. The practice of medical technology. *Sociol Health Illn* 2003;25:97-114 [FREE Full text] [doi: [10.1111/1467-9566.00342](https://doi.org/10.1111/1467-9566.00342)] [Medline: [14498932](https://pubmed.ncbi.nlm.nih.gov/14498932/)]
51. Aarts J, Callen J, Coiera E, Westbrook J. Information technology in health care: socio-technical approaches. *Int J Med Inform* 2010 Jun;79(6):389-390. [doi: [10.1016/j.ijmedinf.2010.03.006](https://doi.org/10.1016/j.ijmedinf.2010.03.006)] [Medline: [20403726](https://pubmed.ncbi.nlm.nih.gov/20403726/)]

52. Emergency Broadband Benefit. Federal Communications Commission. 2021. URL: <https://www.fcc.gov/broadbandbenefit> [accessed 2021-05-18]

Abbreviations

ADD2PCOR: Addressing the Digital Divide to Improve Patient-Centered Outcomes Research

FACETS: Functional Assessment of Comfort Employing Technology Scale

Edited by G Eysenbach; submitted 21.05.21; peer-reviewed by L Welch, K Colorafi, W Buente; comments to author 30.08.21; revised version received 03.09.21; accepted 13.09.21; published 10.11.21.

Please cite as:

Drazich BF, Nyikadzino Y, Gleason KT

A Program to Improve Digital Access and Literacy Among Community Stakeholders: Cohort Study

JMIR Form Res 2021;5(11):e30605

URL: <https://formative.jmir.org/2021/11/e30605>

doi: [10.2196/30605](https://doi.org/10.2196/30605)

PMID: [34757316](https://pubmed.ncbi.nlm.nih.gov/34757316/)

©Brittany F Drazich, Yeukai Nyikadzino, Kelly T Gleason. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.11.2021. 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

Public Perceptions of Diabetes, Healthy Living, and Conversational Agents in Singapore: Needs Assessment

Dhakshenya Ardhithy Dhinakaran¹, BSc (Hons); Thirunavukkarasu Sathish^{2,3}, MBBS, MPH, PhD; Tobias Kowatsch^{4,5,6}, MSc, PhD; Konstadina Griva³, PhD; James Donovan Best¹, MBBS, MD; Lorainne Tudor Car^{1,7}, MSc, MD, PhD

¹Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, Singapore

²Population Health Research Institute, McMaster University, Hamilton, ON, Canada

³Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

⁴Future Health Technologies Programme, Campus for Research Excellence and Technological Enterprise, Singapore-ETH Centre, Singapore, Singapore

⁵Centre for Digital Health Interventions, Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

⁶Centre for Digital Health Interventions, Institute of Technology Management, University of St Gallen, St Gallen, Switzerland

⁷Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

Corresponding Author:

Lorainne Tudor Car, MSc, MD, PhD

Lee Kong Chian School of Medicine

Nanyang Technological University Singapore

Clinical Sciences Building, 11 Mandalay Road

Singapore, 308232

Singapore

Phone: 65 69041258

Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: The incidence of chronic diseases such as type 2 diabetes is increasing in countries worldwide, including Singapore. Health professional–delivered healthy lifestyle interventions have been shown to prevent type 2 diabetes. However, ongoing personalized guidance from health professionals is not feasible or affordable at the population level. Novel digital interventions delivered using mobile technology, such as conversational agents, are a potential alternative for the delivery of healthy lifestyle change behavioral interventions to the public.

Objective: We explored perceptions and experiences of Singaporeans on healthy living, diabetes, and mobile health (mHealth) interventions (apps and conversational agents). This study was conducted to help inform the design and development of a conversational agent focusing on healthy lifestyle changes.

Methods: This qualitative study was conducted in August and September 2019. A total of 20 participants were recruited from relevant healthy living Facebook pages and groups. Semistructured interviews were conducted in person or over the telephone using an interview guide. Interviews were transcribed and analyzed in parallel by 2 researchers using Burnard’s method, a structured approach for thematic content analysis.

Results: The collected data were organized into 4 main themes: use of conversational agents, ubiquity of smartphone apps, understanding of diabetes, and barriers and facilitators to a healthy living in Singapore. Most participants used health-related mobile apps as well as conversational agents unrelated to health care. They provided diverse suggestions for future conversational agent-delivered interventions. Participants also highlighted several knowledge gaps in relation to diabetes and healthy living. Regarding barriers to healthy living, participants mentioned frequent dining out, high stress levels, lack of work-life balance, and lack of free time to engage in physical activity. In contrast, discipline, preplanning, and sticking to a routine were important for enabling a healthy lifestyle.

Conclusions: Participants in this study commonly used mHealth interventions and provided important insights into their knowledge gaps and needs in relation to changes in healthy lifestyle behaviors. Future digital interventions such as conversational agents focusing on healthy lifestyle and diabetes prevention should aim to address the barriers highlighted in our study and motivate individuals to adopt healthy lifestyle behavior.

(*JMIR Form Res* 2021;5(11):e30435) doi:[10.2196/30435](https://doi.org/10.2196/30435)

KEYWORDS

conversational agents; chatbots; diabetes; prediabetes; healthy lifestyle change; mobile phone

Introduction

Background

The growing burden of diabetes is a matter of global concern. Among developed countries, Singapore has the second highest prevalence rate, with 1 in 9 people aged between 18 and 69 years having the condition [1]. Singapore's life expectancy increased from 83.2 in 2010 to 84.8 in 2017; however, when adjusted for time in perfect health (health-adjusted life expectancy), Singapore's life expectancy dropped to 74.2 [2]. The main contributors to this drop are chronic conditions such as diabetes, hypertension, and high cholesterol levels, which are largely due to unhealthy lifestyle and habits [2]. Approximately 440,000 Singaporean residents (aged 18 years and above) were diagnosed with diabetes in 2014, and this number is expected to reach 1 million by 2050 [3].

Prediabetes is a precursor of type 2 diabetes and is becoming increasingly common. Lifestyle interventions, including promotion of physical activity and a healthy diet, delivered by trained health professionals can reduce the incidence of type 2 diabetes in those with prediabetes [4-6]. However, personalized guidance and attention from expert health care professionals are not feasible or affordable on a large scale. A potentially more accessible alternative to in-person support could be digital initiatives for information and intervention delivery, such as conversational agents. Moreover, health programs delivered over the internet have been successful, as exemplified by web-based interventions to reduce smoking [7] and alcohol intake [8] and to improve sexual health [9], cancer screening [10], physical activity [11], and diet [12]. The ubiquity of the internet makes these programs easily accessible to diverse groups. Conversational agents or chatbots are computer programs designed to mimic human-human conversations using either text or speech. Conversational agent use has advantages such as easy access, the possibility of personalization, greater efficiency, bidirectional interactivity, and a chance to build up a working alliance, highlighting the potential to improve patient care [13]. Applications of conversational agents in health care are gaining traction in a number of medical fields and diverse age groups (from children to older adults). Thus far, they have been employed in health care service provision, chronic disease management, patient education and can be delivered via existing messaging apps, individual apps, or even standalone devices [14,15].

People's knowledge of diabetes, prediabetes, and the role and impact of healthy living can affect their lifestyle choices and ultimately health [16]. Before introducing a novel method of digital health delivery to a study population, a needs assessment study is required. This process will identify the challenges faced by the population, the essential components they wish to see, and gauge their likely acceptability of a new digital initiative for their health [17]. Evidence on conversational agent use for healthy lifestyle promotion in different settings, including

Singapore, is currently limited. Singapore, a technologically savvy country, has a high use of mobile phones and messaging apps, as evidenced by its high smartphone penetration rate of 82% in 2021 [18,19]. Singapore's Ministry of Health has highlighted the increasing use of conversational agents in health care to tackle issues such as the rising chronic disease burden within an aging population [20]. This situation makes Singapore an ideal candidate for the implementation and testing of novel mHealth interventions, such as conversational agents [21].

Objectives

To address this need, this study aims to design and evaluate a conversational agent that promotes healthy lifestyle behavior changes for the general population in Singapore. Education and widespread delivery of a healthy lifestyle intervention at the level of the general population can be pivotal in the prevention of diabetes and prediabetes even before individuals get to a stage of high risk. To inform the development of such a conversational agent, we invited members of the public to share their views on and experience of diabetes, prediabetes, healthy living, and digital health interventions, which we report in this study.

Methods

This qualitative study was conducted between August and September 2019. We invited 20 members of the public to participate in the semistructured interviews. The study was approved by the Nanyang Technological University Ethics Committee (IRB-2018-11-032). All participants read the study information sheet before providing written consent. Specific consent was obtained to record the interviews. This study followed the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines [22] ([Multimedia Appendix 1](#)).

Participants and Recruitment

Participants were volunteers initially recruited for a healthy lifestyle change conversational agent pilot feasibility study using a study poster on relevant healthy living Facebook (Facebook Inc) pages and groups ([Figure 1](#)). Participants taking part in this study had to read and sign informed consent, complete an eligibility, baseline, and follow-up questionnaire, in addition to conversing with a Facebook Messenger conversational agent over the 4-week study period. We included adults aged ≥ 21 years who were fluent in English with a Facebook account. The following exclusion criteria were applied:

- individuals with a history of major illness, such as cancer, heart disease, stroke, chronic liver disease, chronic kidney disease, neurodegenerative condition, and hypertension
- individuals with physical disability that would prevent regular physical activity
- pregnant women
- participants younger than 21 years of age
- illiterate or nonwriting individuals (as all questions will be asked in English)

Figure 1. The order of events participants underwent, from (1) recruitment, (2) interviews for needs assessment and subsequent participation in a 4-week feasibility study.



We excluded individuals with a history of major illness, those with physical disabilities, or pregnant women, as advice provided by the conversational agent was not optimized to consider diet and exercise requirements for these populations. In addition, as the content in questionnaires, informed consent, interviews, and conversational agents were all in English, illiterate, and nonwriting individuals were excluded.

We used a purposive sampling strategy to capture variations in ethnicity and age groups from a wider range of perspectives.

Data Collection

Interviews were conducted by a female PhD student (DAD) in designated private meeting rooms at the Lee Kong Chian School of Medicine, NTU, Singapore. An interview guide was used, and interviews were conducted either in person or over the telephone, depending on the volunteer's preference. DAD was provided with sufficient details, resources, and exposure to web-based courses on qualitative research and conducting telephone interviews before study commencement. Field notes were taken during the interviews. The initial interview guide was informed by the literature on the development of digital and conversational agent-delivered health interventions ([Multimedia Appendix 2](#)) [23-26]. The interview was adapted further as the study proceeded, to take account of emerging themes, and each interview ended when we reached saturation of novel topics (eg, for barriers and facilitators, respectively). The interview was an earlier phase in a larger feasibility study, in which all the participants were enrolled ([Figure 1](#)). This study aims to test the feasibility and acceptability of a conversational agent for healthy lifestyle changes in Singapore. Participants were compensated with a digital voucher of SGD \$25–35 (US \$18–25) upon pilot feasibility study completion, and there was no extra incentive for taking part in the telephone interviews. Participation was voluntary.

Basic demographic data were obtained from all participants, including gender, age, ethnicity, marital status, monthly household income, number and age of children, and occupation.

Their use of conversational agents and feedback on these interactions and suggestions for future health applications (eg, tone, media, and direction) were also explored. These insights would unveil their preferences, which would be useful in directing our design and development of conversational agents for this population. Participants were also asked to share the types of messaging and health apps they had experience using and their opinions on them. Information on their use of messaging apps was important to note, as many conversational agents were delivered via these platforms. Public preferences can direct an appropriate delivery platform for these agents. The naming and patterns of use of health apps would indicate

the comfort of the population currently using technology to monitor their health and what applications they find mHealth solutions appropriate for.

Questions were asked about participants' knowledge of diabetes and prediabetes as well as their family history of these conditions. Identifying their knowledge gaps and awareness of being at risk could help in preventing and reducing the risk of developing diabetes and prediabetes. Finally, to best understand their needs on how to promote healthy living, questions were asked about what they thought were the barriers and facilitators of a healthy lifestyle. The focus was on these topics: eating habits, physical activity, stress, and sleep. Participants were asked about their current habits and thoughts on what the ideal advice for these components should be. This would inform the content of future conversational agents educating the public on healthy lifestyle changes.

Both short- and long-answer questions were asked. An example of a short answer question is *Do you use messaging apps such as Facebook Messenger?* An example of a question that required a more comprehensive answer is *Sometimes we do not always eat the way we would like to. Let's talk about some difficulties you experience with eating healthy.*

All in-person interviews were conducted using a portable audio recorder. Telephone interviews were conducted via a speaker phone using the departmental iPhone 6 (Apple Inc) and recorded using an audio recorder placed in close proximity to the speaker phone and the interviewer. The interview lasted approximately 30 minutes. The audio files were transcribed verbatim. All transcripts were checked with the corresponding sound files for accuracy, and they were subsequently corrected for errors and cleaned.

Data Analysis

The data were analyzed using Burnard's method, a structured approach for thematic content analysis established in 1991 and is still used today [27]. First, 2 researchers familiarized themselves with the transcripts by reading them multiple times. Second, the initial codes were proposed. Third, the themes were derived from the codes. Fourth, the 2 authors discussed and combined their themes for comparison. Finally, they reached a consensus on the themes to be used and how to define them. The coding of transcripts was performed using the *review* and *comments* function on a word processor. The final codebook is presented in [Multimedia Appendix 3](#).

Results

Overview

We approached the 60 participants enrolled in the conversational agent feasibility study (via email), looking for 20 volunteers to participate in a telephone interview (Figure 1). We chose this number based on the recommendation that qualitative analysis should be based on data from 1–30 respondents [28]. A total of 20 participants responded and agreed to be interviewed. Seven interviews were conducted in person, and 13 were conducted over the telephone.

Participant Demographics

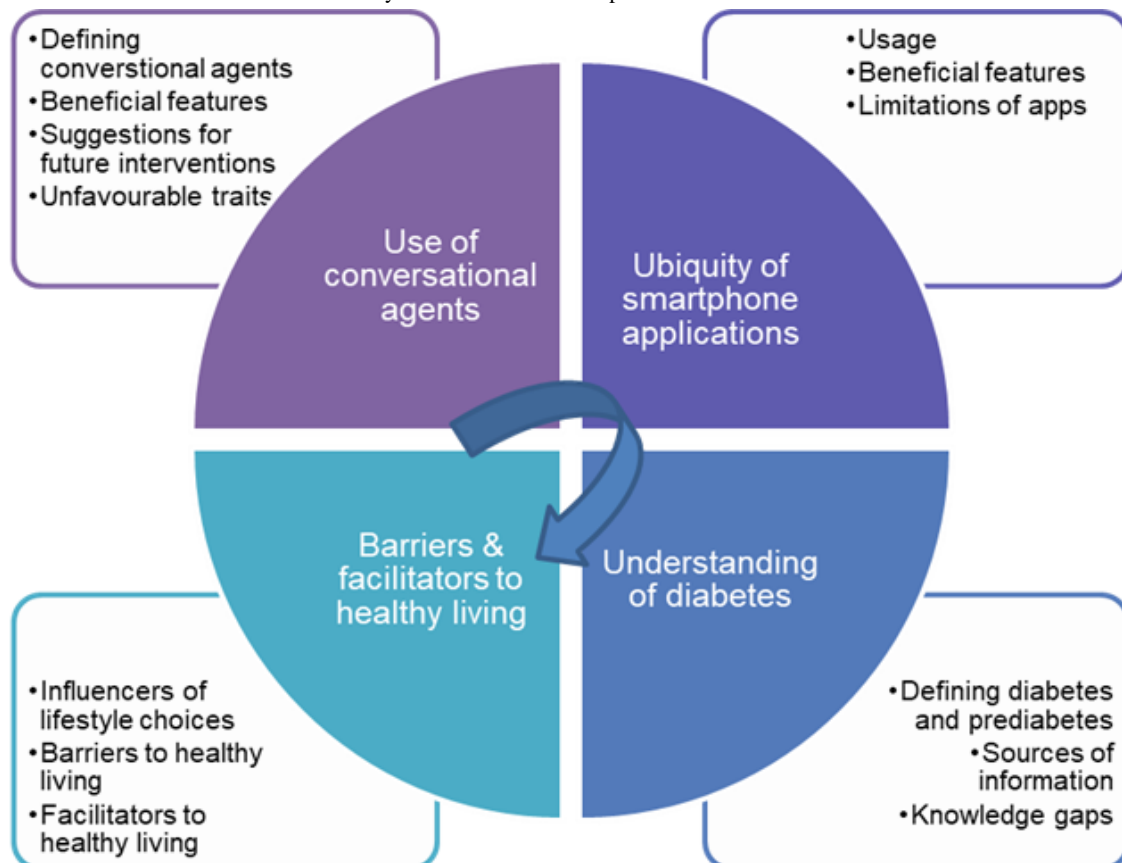
Eighty percent of the participants were female, median age was 30 (in complete years), age range was 23–60 years, and 50% (10/20) were married. Of the 20 participants, 65% (13/20) were working, 25% (5/20) were full-time students (undergraduate or

postgraduate), and 10% (2/20) were unemployed. More than half of the participants (11/20, 55%) were Chinese, 25% (5/20) were Indian, 15% (3/20) were Malay, and 5% (1/20) was White. The monthly household income was SGD \$6000–8000 (US \$4500–6000) for 25% (5/20) of the participants, SGD \$8000–10,000 (US \$6000–7500) for 20% (4/20), SGD \$10,000–15,000 (US \$7500–11,000) for 20% (4/20), SGD \$4000–6000 (US \$3000–4500) for 15% (3/20), less than SGD \$4000 (US \$3000) for 15% (3/20), and more than SGD \$15,000 (US \$11,000) for 5% (1/20). The median monthly household income in Singapore in 2019 was SGD \$9425 (US \$7000) [29].

Thematic Analysis

Four main themes were identified: (1) use of conversational agents, (2) ubiquity of smartphone apps, (3) understanding of diabetes, and (4) barriers and facilitators to a healthy living in Singapore (Figure 2).

Figure 2. Themes and subthemes identified from analysis of interview transcripts.



We present each theme below, along with subthemes and excerpts that highlight different facets of the theme.

Theme 1: Use of Conversational Agents

Definition of Conversational Agents

Many participants were able to provide a definition for what a conversational agent was by mainly using the key words *AI programme* or *virtual call centre* with which one can communicate. Participants were also aware that conversational agents were readily available at all times of the day. Existing applications include car insurance, public transport, web-based

shopping, banking, and airline services. A couple of participants were unfamiliar with this term.

Beneficial Features and Suggestions for Future Interventions

They shared their views on what makes a conversational agent appealing. Conversational agent preferences included accuracy of information, interactivity, visual aids, text, predefined options, a group component, novelty, good reach, incentives, appropriate tone, frequency, duration, and bidirectional communication. Seven participants showed a preference for visual aids (ie,

images and videos) as well as text when communicating with the conversational agent.

Predefined options were preferred over AI if the conversational agent was not sufficiently programmed to deal with more complex questions and requirements. Two participants indicated a liking for peer support in a conversational agent intervention to support their behavior change. Some participants wanted novelty to be introduced in the conversational agent's delivery method to avoid the routine from becoming too monotonous and boring to continue:

But that one was not very good because they cannot... It becomes a pattern. Once they get used to it, then people will try to run away from the pattern. [P15]

Novelty was proposed in the timing of message delivery and message format. For example, some may provide educational information or reminders for healthy behaviors and others, a quick tip for the day.

Conversational agents that had a good reach (ie, were relevant to a wide range of age groups) were most appreciated. Participants mentioned that web-based or digital health platforms such as apps and conversational agents would be well received by the younger population. They indicated a need to present all the relevant information, advice, and reminders on healthy living to older populations who may be less tech savvy by employing simple technology or supplementing with nondigital means. Conversational agents that offered incentives in the form of rewards or discounts were also of interest. Two participants explicitly indicated greater interest in using those apps or conversational agents that, for example, provided discounts for grocery shopping, or coupons for redemption.

Participants had varied views on a suitable frequency and duration of interaction, ranging from daily to weekly. The average suggestion from all 20 participants was a 2-3 days a week frequency, for a duration <20 minutes. Participants were generally indifferent to the tone of a conversational agent but did mention that layman terms should be adopted, and generally, a friendly and informal tone was preferred. Bidirectional communication (where the user can respond and talk to the conversational agent) was preferred over unidirectional communication (where only the conversational agent talks and the user consumes information). Nine participants shared that bidirectional communication would give them the option to ask questions to the conversational agent, clarify doubts, or ask for more information on a topic of interest:

I think two way should be an option. In case like there are any questions or anything you might want to follow up with. Or as well as like... you're interested in something, you can like ask for a contact or they could get you to a website for more information or somewhere that kind of thing. [P05]

Unfavorable Traits

Participants also shared instances in which conversational agent use was unfulfilling. Five participants had negative experiences in the past, where the conversational agents they encountered

were unable to answer their questions or provided irrelevant answers:

When I had questions, they were not answered appropriately. So, I had to resort to calling them. [P09]

Others indicated that sometimes they were unsure of the conversational agent's level of intelligence and realized through the interaction that the agent was not sufficiently artificially intelligent to address their more complex questions. In such cases, they would then prefer a simpler agent (with predefined options) that could fulfill its tasks more seamlessly.

Theme 2: Ubiquity of Smartphone Applications (Specifically Messaging and Health Apps)

Use of Mobile Apps

As conversational agents are frequently embedded in existing messaging apps or implemented as standalone apps, we explored participants' use and opinions of messaging and health apps. All participants mentioned that they used messaging apps, and 17/20 (85%) used health apps. The messaging apps used were WhatsApp (Facebook Inc), Facebook Messenger (Facebook Inc), Telegram (Durov Software), WeChat (Tencent Holdings Limited), and LINE (Naver Corp). These were used mainly for communicating and socializing, sometimes in large groups. Some use it as a means of sharing information for educational and professional purposes.

A variety of health apps were mentioned with *Healthy 365* and step trackers were the most frequently used. Other apps were MyFitnessPal, AIA vitality, Fitbit, Samsung Health, Health Promotion Board app, Nike Run Club, Chronometer (monitor nutrition), and Stride (habit tracker).

Beneficial Features

Specific smartphone app features and functions that make apps more favorable for use were discussed. Eight participants attributed the beneficial features of these apps to fulfilling a necessity, ease of use, engagement (promoted by the availability of stickers), and their multifunctionality, such as texting, calling, and video calling.

Something else participants shared which promotes smartphone app use, is the provision of incentives:

And then for the Healthy 365 app by HPB (Health Promotion Board is a Singapore Government organisation), I use it to redeem the rewards and participate in their promotional activities and programs. [P17]

Limitations of Apps

Two participants reported that sometimes the content and frequency of messages could be too overwhelming to deal with. The accuracy of the app content was sometimes questioned by participants (2/20, 10%), with different apps reporting different numbers for the same measure.

Theme 3: Understanding of Diabetes

We explored the participants' understanding and knowledge of diabetes, prediabetes, and the sources of information that they

used to obtain relevant information. Understanding their current level of knowledge and what information they need more exposure to could be incorporated into a healthy lifestyle conversational agent to prevent diabetes and prediabetes in the general population.

Diabetes

All participants were able to provide a definition of diabetes, some more comprehensive than others. Five individuals mentioned insulin or the inability of the body to control their blood sugar level, which is too high. Four participants stated that diabetes requires one to reduce sugar intake.

They were also aware that there may be a need to take medication and that there is no cure for diabetes. Participants shared the impact of diabetes on their health. Five participants were aware that diabetes can affect other parts of the body and lead to other illnesses, including kidney failure, stroke, high blood pressure, and limb amputations.

Regarding impact on one's life, participants mentioned that diabetes affects families, ability to exercise, and breathing, and can lead to higher medical expenses:

Having good health to be drained away. And then, how does this affect families concerned...I was told that people with diabetes may not be able to exercise actively. Have to change the type of exercise that they do. Sometimes breathing will be affected, also. [P14]

Prediabetes

Fifteen participants were able to define prediabetes, stating that it was a condition where someone is at a borderline risk of developing type 2 diabetes. Two individuals reported that prediabetes was reversible:

Basically, you are at risk of diabetes if you don't change the way you live your life, or your diet. [P20]

Other participants were very unfamiliar with the term *prediabetes* (5/20, 25%) and acknowledged that diabetes was more commonly talked about:

I always saw this like pretype or something number 2 or something like that so I don't really understand it as well [P08]

Sources of Information

The sources of information that they used to learn more about diabetes varied. Some shared that their personal experiences with diabetic family members helped them to be educated on diabetes and its implications (2/20, 10%):

And you know when she (mother) was around, there's a diabetes society or something we used to go and visit, attend talks and yea, somewhere in(...) it's quite near my place actually but after she passed away uh I'm not involved in anything anymore [P01]

Others mentioned the government's role in educating the public on diabetes, prediabetes, and healthy living and how the government can further influence the public in the near future (3/20, 15%):

But from what I used to experience, the government does have certain policies acting through MOE (Ministry of Education). For example, last time we used to have no fried food in canteen. I suppose they must be doing something with school canteens because of the diabetes. [P17]

In addition, participants mentioned that a dearth of sufficient information was readily available to them, causing a barrier to awareness on healthy living (6/20, 30%). Some participants felt that the older generation may have less information because of lower levels of technological competency in searching for information on the web.

Participants proposed additional means to improve information availability and population awareness. Some of these methods include advertisements on the television and internet, as well as a display of the number of calories for each food item at hawker centers (food courts with affordable local meals):

and also there could be more awareness like you know when we go to restaurant or maybe hawkers more information could share like the calories. Because yea I know some hawker centres they do show but it didn't exactly show the number of calories for all the food. [P08]

Knowledge Gaps

Participants identified some areas where they thought societal understanding was low and should be addressed in the future. One such example was more information on physical and psychological changes for a healthier lifestyle:

What the individual lifestyle changes that they can make or consider to make. [P12]

Yes. Definitely. I think when you are more stressful, then your sugar goes up too. [P20]

Mental health. [P15]

They made specific references to education on the importance of healthy eating and physical activity. One participant felt that instead of learning more, it was more important to apply what is already known about diabetes.

Further information on symptoms, prevention, risk factors, and awareness of being at risk were of interest. In addition, there is a need for more information on the implications of diabetes, its management, and sources of support:

Maybe even things like prognosis if you are diagnosed, how are you going to manage it. If they need help, who can they look for. [P11]

Theme 4: Barriers and Facilitators to a Healthy Living in Singapore

Influencers of One's Lifestyle Choices

Participants shared their existing lifestyle habits and the factors that enable or hinder healthy living. These barriers are helpful indicators of what a future conversational agent can aim to address when administering advice, education, and support for healthy lifestyle changes.

They first discussed the relationship between their existing lifestyle choices and how this relates to their perceptions of current body weight. Some participants reported dissatisfaction with their current weight. They perceived themselves as either too heavy (9/20, 45%) or underweight (2/20, 10%).

Participants also highlighted the existing factors which drive their lifestyle choices.

Convenience

Convenience was a big factor. This was relevant regarding food available in the close vicinity, as well as the convenience of having a gym or group exercise class easily accessible to them.

Cost

Cheaper food was preferred because of its affordability, over more expensive, healthier options. Participants also indicated a desire to engage in group exercise sessions, but only if they were affordable:

The biggest problem we have here is the price of healthy food. It's just infinitely more costly. For example, some salad can cost you six, SGD \$7 (US \$5), whereas chicken rice can cost you SGD \$3 (US \$2). You cannot possibly eat salad every day even if you are earning okay. But if you are one of those that are low income, definitely you will go for three-dollar chicken rice every day. And in the long run, it just causes a lot of problems. [P17]

Maybe I'd join Zumba. I like that kind of thing but I'm thinking about affordability and the convenience. [P18]

Personal Preferences

Some participants chose to eat food based on their personal preferences (5/20, 25%). Similarly, they chose to engage in physical activities that they enjoyed or suited their schedules:

Because I know that I'm not really hectic with the exercising, so I would tend to think that my house is three bus stops away from the interchange, so I will tell myself let's just walk from the MRT back home, so at least it's some walking activity for me. [P18]

Social Activities

Some participants chose what to eat based on the kinds of social activities they engaged in and the food available there (3/20, 15%). Furthermore, their social life was influenced by the amount of exercise they could do as more time spent socializing left less time for physical activity (2/20, 10%).

Advertising and Marketing

Marketing of food was a factor influencing the kind of food they decided to eat in a day (2/20, 10%):

It just depends on what I feel like eating or maybe a bit of marketing as well. For example, if McDonald's has a new menu, then I might just go for it [P17]

Nutritional Value

Nutritional value of food determined whether participants chose to eat it:

If I look at the menu item and then it looks like this is 600 calories, which is top of the list out of the whole list of things that I'm looking at, I would think twice about eating it

It does have some sort of impact knowing that you are intentionally consuming so much. You will trigger some automatic response to self-regulate. [P17]

State of Mind

Some participants mentioned that their mood or state of mind largely determined the food they decided to consume (4/20, 20%). For example, when upset, stressed, or in a bad mood, they gravitated toward food that was more comfortable than healthy. Likewise, their state of mind also influenced their physical activity (3/20, 15%):

like after a long day of work, I'm not like oh I'm going to go to the gym you know? Like nah I just want to stay at home and relax. [P05]

Barriers to Healthy Living

Participants shared some barriers to healthy living.

Dining Out

Having to dine out on a regular basis meant that participants had no control over the nutritional value of their food:

Sometimes if you like to go to a restaurant and you hope that there's not so much MSG (monosodium glutamate-flavour enhancer) its really beyond your control to tell the chef hey I don't want more... I don't want the MSG or I want less salt. I mean you can tell them less salt. But pre-prepared stuff usually its hard. [P06]

Lack of Appeal

Two participants mentioned that healthy foods tend to have an unappealing taste and texture, which discouraged its consumption:

Eating healthy, sometimes, they say is very bland, very tasteless. No fried is something without aroma. So, eating healthy sometimes lack of the aroma. And then they say it's either very too liquid, or so tasteless, that is the thought they have. [P14]

Health Condition

One participant shared that a health condition (brain tumor) brought about compulsive eating habits, which caused her to deviate from a healthier diet.

Fear of Injury

Two participants also mentioned a fear of injury or internal damage to their bodies, which may be caused by physical activity, preventing them from exercising. Hence, this fear became a barrier to a more active lifestyle:

but there is also a fear that at a certain age there is some wear and tear in the body. [P14]

Lack of Free Time

Long working hours were associated with less free time and, hence, less frequent exercise. This lack of work-life balance also exacerbates stress:

whether I have time to dock off for a run or not. It really depends because most of the time, I'm actually working quite late.... I think for myself would definitely be work stress.... So myself sometimes I work 15 to 16 hours a day. [P08]

Specific Events

Specific events, such as unexpected incidents, can disrupt an individual's equilibrium and cause stress. Similarly, events such as exams can negatively impact sleep schedules. Changes in weather, such as a hazy or wet period in Singapore, can prevent individuals from being physically active if there is a preference to exercise outdoors.

Lack of Knowledge

Lack of knowledge was mentioned by several participants (11/20, 55%) as a barrier to healthy living. For example, some participants had trouble going to bed early and did not know how to fix this problem. Others were unsure about how to manage their stress, and hence, the problem persists until they find or learn a solution:

I'm having a bit of difficulty here because I'm experiencing a new sort of stress after coming back from my maternity leave. So, that's why I don't really know what to tell you because I think I haven't really learned how to manage the stress. [P13]

Facilitators of Healthy Living

Participants shared some factors which can help to promote healthy living.

Early Intervention and Knowledge Building

Building knowledge on healthy living from a young age was stated as a promoter. Building more as an adult on existing knowledge of the importance of healthy behaviors could positively influence individuals.

Reminders

Although the knowledge may already be there, reminders were cited as a way to reiterate healthy eating behaviors. Participants also requested education on calorie counting, which would promote healthier eating:

So, maybe when it comes to before lunchtime, two hours, tell them this is a suggested meal, then it's a guideline. Because everybody like to hold their phone, and then the message pops up, and they say, oh, yes, why not, I go and try. [P14]

Incorporate Activity Into Existing Routines

Sourcing out alternative methods of reaching the recommended weekly level of physical activity and how to be more active in general were found to be helpful. These came in the form of efficient exercising or reducing the amount of time spent sitting after a meal:

So, yes, I only exercise 15 minutes each time. But I try to keep it more vigorous. So, yes, I do HIIT kind of training. So I do that about four or five times a week maybe. I think that there's more yield in doing vigorous exercise over moderate exercise. So in terms of the efficiency, the amount of time you need to spend, so I try to do more of HIIT. [P20]

Empowering Oneself

Empowering individuals to provide services for themselves as much as possible—for example, making homemade food was suggested to be an effective way of eating healthier. Engaging in exercise at home and setting up their own workouts were also mentioned. Similarly, healthy self-management of stress in the form of therapeutic activities, for example, was cited as a good way to maintain a healthier lifestyle:

So I have a treadmill at home and I use that. Like I walk as I'm doing some work on my laptop. So I try to incorporate that. Or I might do maybe like 10/15 minutes of yoga or cardio or something just around the house yea wherever I can. [P05]

Group Support

It was proposed by 4 participants that group support, which evokes a collaborative spirit could encourage more physical activity:

If I did it with other people, like with friends and stuff yea, you could make it like a group effort. Then it'll be easier because then there are other people to monitor you, just for yourself. [P05]

Internal Motivators

Internal motivators, such as wanting to be healthy in the present and future, encouraged individuals to take ownership and eat healthily.

Discipline

Enforcing some discipline with regard to eating on time, drinking enough water, and planning meals for the week were all suggestions for promoting a healthier diet:

But personally, I do my own meal prep actually. So, I cook for three days, so three of the lunches is covered for me. [P11]

Having good discipline and planning slots for physical activity were mentioned as effective methods for developing a more active lifestyle. Discipline in developing and sticking to a bedtime routine that involves winding down, relaxing, and reducing screen time as well as cutting down caffeine intake were proposed by participants to improve sleep habits. Recommendations for disciplining food intake through moderation involved a balance of healthy and unhealthy foods, controlling portion sizes (overall quantity and food pyramid recommended portions), and restricting unhealthy food items. These choices were cited as effective promoters for healthy living:

For example, bubble tea right, everybody likes bubble tea, so to me I will limit myself to buy that maybe once

every two months, take that as a reward for myself, yes, that's what I'd do. [P18]

Moderation

One participant recommended establishing moderation at work to have a more work-life balance to reduce stress levels:

And I think having a good balance, life, social life, helps as well, that it's not just work. [P11]

Participants also mentioned compensating for sleep over the weekends to moderate overall sleep quantity.

Discussion

Principal Findings

To our knowledge, this is the first qualitative study to investigate the public's perception of conversational agents for a healthy lifestyle change intervention in Singapore.

Participants' perceptions of conversational agents were discussed, where they offered definitions, usage experiences, and preferential features. These are pivotal in informing the design elements for a feasible and acceptable conversational agent initiative in a health care setting in Singapore. Participants shared their views on the ubiquity of smartphone apps, whereby messaging apps were a necessity for daily communication, social, educational, and professional purposes. This is an important point, as conversational agents can be embedded in messaging app platforms that are most familiar to the target population for easy adoption. Health apps were used for a variety of purposes such as step tracking, health tracking (nutrition, water intake, and habit tracking), or pairing with a Fitbit watch. The fairly diverse patterns of use indicate an inclination to adopt mHealth solutions for numerous purposes. Furthermore, participants' views on the beneficial and limiting features of apps and conversational agents are highly relevant and should be considered when developing a health conversational agent for this population. Participants shared their existing knowledge on diabetes, prediabetes, and their existing sources of information. In addition, they offered their views on what constitutes healthy living with a focus on diet, exercise, sleep, and stress as well as thoughts on the barriers and facilitators of healthy living in Singapore. The culture of eating out in Singapore was said to be fueled by affordable hawker centers, fast food, and fast-paced life, which does not allow much time for food preparation at home. Frequent dining out, high stress levels, lack of work-life balance, and lack of free time to engage in physical activity were some of the common complaints hindering a healthier lifestyle. In addition, work-related stress has been implicated in poor sleep quality and quantity. Participants proposed that these barriers could be tackled with some self-discipline, preplanning, and sticking to a routine for healthier living patterns. Given the ubiquity of smartphones, the avid use of messaging apps and an inclination to use mHealth initiatives and awareness of these barriers and facilitators for healthy living could be communicated to the population via a conversational agent.

Comparisons With Existing Literature

Lim et al [30] explored the barriers and facilitators of healthy eating in Singapore in 2019. This study noted that components that encouraged healthy eating included self-discipline, fear of disease complications, education by a health professional and mass media, and health promotion campaigns [30]. Participants in our study also mentioned self-discipline as a strong facilitator and acknowledged the presence of health promotion campaigns but did not cite these as influencers of their lifestyle choices. Furthermore, fear of disease complications and education by health professionals was possibly more relevant in Lim et al [30], as participants were recruited from polyclinics and were already diagnosed with prediabetes. This also points out the need for a more hybrid approach when it comes to more high-risk populations (such as those with prediabetes), as HCP involvement would be necessary in addition to a conversational agent. Furthermore, the lack of skills to prepare and choose healthy foods was shared as a difficulty in healthy eating [30]. This barrier was also noted as a limitation to a healthy diet in our study. Participants added that a lack of time to do grocery shopping, prepare and cook meals was due to their hectic schedules, making dining out a more convenient, but often less healthy, compromise.

Another study noted that healthy eating and physical activity were the main preferred components for health education and communication in patients with prediabetes in Singapore [31]. A similar outcome was implicated in our study, but we also considered the components of sleep and stress for a more holistic view of healthy living. Other necessities for health communication were risk and prevention of diabetes [31]. These were also included in the qualitative analysis. In addition, participants shared an interest in information on where to go for help if they were at risk or diagnosed with prediabetes or diabetes.

Other existing studies have investigated novel conversational agent interventions, with some reporting on the acceptability and usability of these agents. However, an extensive needs assessment analysis, such as the one presented here, for a conversational agent intervention has not been performed so far in the Singaporean population. A scoping review of conversational agents showed that qualitative data were presented in some studies to show the acceptability and satisfaction of these interventions, namely, participants' opinions of an already developed intervention [14]. These parameters were often reported with Likert scale ratings, but not comprehensively or thematically [32,33].

In a study on behavior change in overweight adolescents, high compliance to conversational agent intervention was attributed to a rewarding game system [34]. This finding aligns with the suggestions from participants in this study that incentives will be well received and will contribute to their willingness to use the conversational agent. Other studies have examined the use of conversational agents for chronic conditions, including diabetes, and reported a preference for the features of conversational agents that allowed for personalization [35]. Personalization was also noted as an advantageous trait in our study participants.

Ta et al [36] discussed the type of support provided by a popular chatbot available on app stores, *Replika*, with generic capabilities such as stress management and social support. Based on the analysis of public user reviews and a survey conducted with participants recruited via social media in the United States, companionship was the greatest form of social support users identified. They attributed this to *Replika*'s ability to use a range of message types (eg, text and images) and appear human-like. Similarly, interviewees in our needs assessment supported the use of visual aids in addition to text. Although they did not indicate specific personality preferences or the need for the conversational agent to be human-like, a general inclination toward a friendly and informal tone was identified.

Studies on physicians' perceptions of conversational agents have indicated their apprehension toward agents that are unable to comprehend the emotional state of vulnerable patients and have dissuaded their use in cases where expert medical knowledge and skills are required [37]. However, they have supported the use of conversational agents to support, motivate, and coach patients. In the same way, we aim to develop a conversational agent to support a healthy lifestyle change by understanding the needs and preferences of prospective users via this needs assessment.

Implications for Future Work

Future digital interventions aimed at encouraging diabetes prevention in the general population to promote and prolong healthy living can incorporate the content-and delivery-specific options highlighted in [Textbox 1](#).

These suggestions include improving knowledge in areas directly relevant to diabetes prevention, building up an awareness of risk factors, symptoms, and prevention methods, as well as the complications and consequences of diabetes and prediabetes. Future applications can also explore expanding actionable advice on how to apply the knowledge participants acquired to healthy living.

It would be beneficial to help users improve their skills related to food choices, calorie counting, and stress management. Advice on healthy eating should be mindful of the Asian diet that most Singaporeans are familiar with and have a preference for. In relation to calorie counting, the usefulness of sharing information on the number of calories in frequently consumed food items in Singapore hawker centers should be examined. Stress management techniques relevant to establishing a work-life balance would potentially be beneficial, as this is a common complaint in Singapore, as disclosed by our participants. In addition, teaching participants time management

skills may help to reduce overall stress, providing time and motivation for physical activity, meal preparation, and a bedtime routine. Again, the outcomes of this intervention would need to be evaluated.

Options to reach the weekly recommendation for physical activity, such as increased efficiency, safe exercise options, or suggestions to see a doctor for more personalized advice can be made via a digital conversational agent. Regarding group exercise, cheaper options or even web-based options may be listed. This approach comes with the added benefits of convenience (no traveling) and affordability. Recommendations on alternative methods of exercising, which do not require participants to be outdoors, could be suggested to overcome barriers related to the weather, such as home workouts. A conversational agent could also reinforce the existing promoters of physical activity, for example, reminding them of the feelings of contentment they will experience.

Participants in our study also provided informative recommendations for the delivery of future digital health interventions, focusing on healthy lifestyle behavior changes. App or conversational recommendations are proposed. These factors have enhanced user experience in the past. They also shared the components they wish to see and would enjoy using in future applications.

Future developments should investigate personalizing the conversational agent to the user's preferences (timings, advice, etc), enabling easy use, making the exchange engaging, and introducing some novelty to boost user experience. The tone, personality, and language used by the conversational agent should also be adapted to suit the target population (eg, being friendly and using language with a level of complexity familiar to the target population). Some additional considerations, which could be included, are incentives, such as rewards or discounts for groceries, and group support, as participants indicated a preference for these.

On the basis of this study, we are developing a rule-based conversational agent pilot intervention that incorporates recommendations from the participants, including educational information on the definitions and implications of diabetes and prediabetes, as well as actionable advice on healthy eating, physical activity, stress management, and healthy sleep patterns. All the information to be included in the intervention will be evidence-based. We will also incorporate visual aids to supplement the text, and preprogrammed options will be provided for the user to choose from, to introduce some degree of bidirectional communication.

Textbox 1. Recommendations on content and delivery for future conversational agent interventions for healthy lifestyle behavior change.

Content and delivery for future conversational agent interventions

- Content
 - Furthering knowledge on:
 - Implications of diabetes and prediabetes
 - Awareness of being at risk
 - Symptoms and prevention methods
 - Actionable advice on:
 - How to apply the knowledge gained on healthy living
 - Further development of skills relating to:
 - Eating healthily
 - Calorie counting
 - Options available for meeting the weekly recommendation for physical activity
 - Stress management
 - Time management
 - Type of content:
 - Novel—up to date
 - Trustworthy—evidence based
- Delivery
 - Personalization of content
 - Advice on how to incorporate exercise into one's specific schedule, availability of facilities and preferences
 - Habit formation – how to discipline your healthy eating, drinking water, physical activity and bedtime routine
 - Suitable frequency and duration
 - Ask participants for their preference
 - Offer options for participants to choose from
 - Easy use
 - Predefined options for better efficiency if AI is not yet maximized
 - Engaging
 - Use stickers and visual aids
 - Make the conversation bidirectional so the user can be part of the exchange and can ask questions if necessary
 - Relevant to target population
 - Tone, personality and language used by the conversational agent to be easily understood, relevant and appropriately tailored for the target population
 - *Additional considerations to keep users keen and engaged in the interventions
 - Incentives, such as rewards and vouchers
 - Group support, for example, chat channel for participants to communicate with others using the same conversational agent to share advice and motivate each other

Strengths and Limitations

In our study, we recruited and interviewed a sample of 20 participants who were ethnically diverse and covered a wide age range (23-60 years). The collected data were coded in parallel by 2 researchers, with the development and application of a common coding framework through a series of discussions. Our findings may also be applicable to other high-income or technologically savvy Asian countries. Although 20 participants were adequate for this analysis, we would aim for a larger sample size in future studies for greater validity and generalizability.

We extrapolated the focus of this study from healthy living to diabetes and prediabetes prevention. To this regard, we managed to provide suggestions on healthy living not only in the context of wellness but also as a means of preventing conditions of great burden in Singapore—diabetes and prediabetes.

There is a possibility of recruitment bias. The participants were all from Facebook groups and pages with an interest in healthy living. Their opinions may be somewhat skewed as they already came with some knowledge and interest in a healthy lifestyle and some technological competency. In addition, despite our

purposive sampling, 16/20 (80%) of our participants were female because of the interest shown by the study volunteers. Both of these reduce the generalizability of the findings as they may not be entirely representative of the Singaporean population at large.

Conclusions

The participants provided valuable insights on their existing knowledge and sources of information on healthy living and diabetes. They also shared their current patterns of use of messaging apps, health apps, and their views on potential conversational agents, which have not yet been used for health care purposes in Singapore.

In addition, they offered opinions about the importance of healthy living and diabetes prevention, likely to be reflective of a significant part of the at-risk population. Preferences for conversational agents and use of smartphone app were also discussed. Finally, they shared views on barriers and facilitators of healthy living. Our findings can be used to inform the development of future conversational agent interventions and similar mHealth initiatives that target healthy lifestyle behavior changes.

Acknowledgments

This research was supported by the Ageing Research Institute for Society and Education (ARISE), Nanyang Technological University, Singapore. It was also supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) program.

Authors' Contributions

LTC conceived the idea of this study. LTC and DAD developed and prepared the resources. DAD conducted all the interviews. LTC and DAD analyzed the data and interpreted the results with input from KG. DAD, LTC, ST, KG, TK, and JDB. They were involved in writing the paper and approved the submitted version.

Conflicts of Interest

TK is affiliated with the Centre for Digital Health Interventions, a joint initiative of the Department of Management, Technology, and Economics at ETH Zürich and the Institute of Technology Management at the University of St Gallen, which is funded in part by the Swiss health insurer CSS. TK is also a cofounder of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways. Other authors declared that they had no conflicts of interest.

Multimedia Appendix 1

Consolidated Criteria for Reporting Qualitative Research (COREQ): 32-item checklist.

[[DOCX File, 21 KB - formative_v5i11e30435_app1.docx](#)]

Multimedia Appendix 2

Interview guide.

[[DOCX File, 30 KB - formative_v5i11e30435_app2.docx](#)]

Multimedia Appendix 3

Codebook from qualitative analysis.

[[DOCX File, 36 KB - formative_v5i11e30435_app3.docx](#)]

References

1. Jessie Lim. Singapore is No. 2 nation with most diabetics: 5 things about diabetes. The Straits Times. 2016 Apr 08. URL: <https://www.straitstimes.com/singapore/singapore-is-no-2-nation-with-most-diabetics-5-things-about-diabetes>

2. Mothership. 2019. URL: <https://mothership.sg/2019/03/singaporean-life-expectancy-diabetes-hypertension/> [accessed 2021-10-06]
3. 10 Things You Need to Know about Diabetes. Health Hub. URL: <https://www.healthhub.sg/live-healthy/1336/10-things-you-need-to-know-about-diabetes> [accessed 2021-10-14]
4. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001 May 03;344(18):1343-1350. [doi: [10.1056/NEJM200105033441801](https://doi.org/10.1056/NEJM200105033441801)] [Medline: [11333990](https://pubmed.ncbi.nlm.nih.gov/11333990/)]
5. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002 Feb 07;346(6):393-403 [FREE Full text] [doi: [10.1056/NEJMoa012512](https://doi.org/10.1056/NEJMoa012512)] [Medline: [11832527](https://pubmed.ncbi.nlm.nih.gov/11832527/)]
6. Ramachandran A, Snehalatha C, Mary S, Mukesh B, Bhaskar AD, Vijay V, Indian Diabetes Prevention Programme (IDPP). The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). *Diabetologia* 2006 Feb;49(2):289-297. [doi: [10.1007/s00125-005-0097-z](https://doi.org/10.1007/s00125-005-0097-z)] [Medline: [16391903](https://pubmed.ncbi.nlm.nih.gov/16391903/)]
7. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. *Addiction* 2009 Nov;104(11):1792-1804. [doi: [10.1111/j.1360-0443.2009.02710.x](https://doi.org/10.1111/j.1360-0443.2009.02710.x)] [Medline: [19832783](https://pubmed.ncbi.nlm.nih.gov/19832783/)]
8. Khadjesari Z, Murray E, Hewitt C, Hartley S, Godfrey C. Can stand-alone computer-based interventions reduce alcohol consumption? A systematic review. *Addiction* 2011 Feb;106(2):267-282. [doi: [10.1111/j.1360-0443.2010.03214.x](https://doi.org/10.1111/j.1360-0443.2010.03214.x)] [Medline: [21083832](https://pubmed.ncbi.nlm.nih.gov/21083832/)]
9. Bailey J, Murray E, Rait G, Mercer C, Morris R, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. *Cochrane Database Syst Rev* 2010 Sep 08(9):CD006483. [doi: [10.1002/14651858.CD006483.pub2](https://doi.org/10.1002/14651858.CD006483.pub2)] [Medline: [20824850](https://pubmed.ncbi.nlm.nih.gov/20824850/)]
10. Lee EE, Brecht M, Park H, Lee J, Oh KM. Web-based study for improving mammography among Korean American women. *J Cancer Educ* 2017 Jun;32(2):257-263. [doi: [10.1007/s13187-015-0920-2](https://doi.org/10.1007/s13187-015-0920-2)] [Medline: [26423058](https://pubmed.ncbi.nlm.nih.gov/26423058/)]
11. Jennings CA, Vandelanotte C, Caperchione CM, Mummery WK. Effectiveness of a web-based physical activity intervention for adults with type 2 diabetes-a randomised controlled trial. *Prev Med* 2014 Mar;60:33-40. [doi: [10.1016/j.ypmed.2013.12.011](https://doi.org/10.1016/j.ypmed.2013.12.011)] [Medline: [24345601](https://pubmed.ncbi.nlm.nih.gov/24345601/)]
12. Harris J, Felix L, Miners A, Murray E, Michie S, Ferguson E, et al. Adaptive e-learning to improve dietary behaviour: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2011 Oct;15(37):1-160 [FREE Full text] [doi: [10.3310/hta15370](https://doi.org/10.3310/hta15370)] [Medline: [22030014](https://pubmed.ncbi.nlm.nih.gov/22030014/)]
13. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Educ Couns* 2005 Oct;59(1):21-30. [doi: [10.1016/j.pec.2004.09.008](https://doi.org/10.1016/j.pec.2004.09.008)] [Medline: [16198215](https://pubmed.ncbi.nlm.nih.gov/16198215/)]
14. Tudor Car L, Dhinakaran DA, Kyaw BM, Kowatsch T, Joty S, Theng Y, et al. Conversational agents in health care: scoping review and conceptual analysis. *J Med Internet Res* 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: [10.2196/17158](https://doi.org/10.2196/17158)] [Medline: [32763886](https://pubmed.ncbi.nlm.nih.gov/32763886/)]
15. Schachner T, Keller R, V Wangenheim F. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. *J Med Internet Res* 2020 Sep 14;22(9):e20701 [FREE Full text] [doi: [10.2196/20701](https://doi.org/10.2196/20701)] [Medline: [32924957](https://pubmed.ncbi.nlm.nih.gov/32924957/)]
16. Ang BW, Tan MY, Goh CM, Rahardja S, Lim BY, Chiew W, et al. Impact of knowledge and attitudes on lifestyle practices in preventing type 2 diabetes mellitus. *Ann Acad Med Singap* 2019 Aug;48(8):247-263 [FREE Full text] [Medline: [31628744](https://pubmed.ncbi.nlm.nih.gov/31628744/)]
17. Kulandaivelu Y, Lalloo C, Ward R, Zempsky WT, Kirby-Allen M, Breakey VR, et al. Exploring the needs of adolescents with sickle cell disease to inform a digital self-management and transitional care program: qualitative study. *JMIR Pediatr Parent* 2018 Sep 25;1(2):e11058 [FREE Full text] [doi: [10.2196/11058](https://doi.org/10.2196/11058)] [Medline: [31518307](https://pubmed.ncbi.nlm.nih.gov/31518307/)]
18. Singapore tops mobile app use and downloads: App Annie. CNA. 2019. URL: <https://www.channelnewsasia.com/news/technology/singapore-tops-mobile-app-use-and-downloads-app-annie-11157124> [accessed 2021-10-06]
19. Smartphone penetration rate as share of the population in Singapore from 2010 to 2020 and a forecast up to 2025. Statista. URL: <https://tinyurl.com/whrjbk6k> [accessed 2021-10-06]
20. Baker J. Chatbots, robotics to help ease manpower crunch in healthcare sector. Channel News Asia. 2018. URL: <https://tinyurl.com/6vzy58vu> [accessed 2021-10-06]
21. Erickson G, Rothberg H. Chapter 8 - Knowledge and knowledge-related assets: design for optimal application and impact. In: *Successes and Failures of Knowledge Management*. Boston: Morgan Kaufmann; 2016.
22. 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/)]
23. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. *J Med Internet Res* 2015 Oct 23;17(10):e240 [FREE Full text] [doi: [10.2196/jmir.4897](https://doi.org/10.2196/jmir.4897)] [Medline: [26499966](https://pubmed.ncbi.nlm.nih.gov/26499966/)]

24. Fukuoka Y, Gay CL, Joiner KL, Vittinghoff E. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. *Am J Prev Med* 2015 Aug;49(2):223-237 [FREE Full text] [doi: [10.1016/j.amepre.2015.01.003](https://doi.org/10.1016/j.amepre.2015.01.003)] [Medline: [26033349](https://pubmed.ncbi.nlm.nih.gov/26033349/)]
25. Wong CK, Fung CS, Siu S, Lo YY, Wong K, Fong DY, et al. A short message service (SMS) intervention to prevent diabetes in Chinese professional drivers with pre-diabetes: a pilot single-blinded randomized controlled trial. *Diabetes Res Clin Pract* 2013 Dec;102(3):158-166. [doi: [10.1016/j.diabres.2013.10.002](https://doi.org/10.1016/j.diabres.2013.10.002)] [Medline: [24466598](https://pubmed.ncbi.nlm.nih.gov/24466598/)]
26. Ramachandran A, Snehalatha C, Ram J, Selvam S, Simon M, Nanditha A, et al. Effectiveness of mobile phone messaging in prevention of type 2 diabetes by lifestyle modification in men in India: a prospective, parallel-group, randomised controlled trial. *Lancet Diabetes Endocrinol* 2013 Nov;1(3):191-198. [doi: [10.1016/S2213-8587\(13\)70067-6](https://doi.org/10.1016/S2213-8587(13)70067-6)] [Medline: [24622367](https://pubmed.ncbi.nlm.nih.gov/24622367/)]
27. Burnard P. A method of analysing interview transcripts in qualitative research. *Nurse Educ Today* 1991 Dec;11(6):461-466. [doi: [10.1016/0260-6917\(91\)90009-y](https://doi.org/10.1016/0260-6917(91)90009-y)] [Medline: [1775125](https://pubmed.ncbi.nlm.nih.gov/1775125/)]
28. Fridlund B, Hildingh C. *Qualitative Research Methods in the Service of Health. -: Professional Pub Service; 2021.*
29. Are you earning enough? Singapore's average household income revealed. Standard Chartered. 2019. URL: https://www.sc.com/sg/stories/financial-tips/are-you-earning-enough-singapores-average-household-income-revealed/#_ftn2 [accessed 2021-10-06]
30. Lim RB, Wee WK, For WC, Ananthanarayanan JA, Soh YH, Goh LM, et al. Correlates, facilitators and barriers of healthy eating among primary care patients with prediabetes in Singapore-a mixed methods approach. *Nutrients* 2019 May 06;11(5):1014 [FREE Full text] [doi: [10.3390/nu11051014](https://doi.org/10.3390/nu11051014)] [Medline: [31064063](https://pubmed.ncbi.nlm.nih.gov/31064063/)]
31. Lim RB, Wee WK, For WC, Ananthanarayanan JA, Soh YH, Goh LM, et al. Health education and communication needs among primary care patients with prediabetes in Singapore: a mixed methods approach. *Prim Care Diabetes* 2020 Jun;14(3):254-264. [doi: [10.1016/j.pcd.2019.08.008](https://doi.org/10.1016/j.pcd.2019.08.008)] [Medline: [31558372](https://pubmed.ncbi.nlm.nih.gov/31558372/)]
32. Comendador BE, Francisco BM, Medenilla JS, Nacion SM, Serac TB. Pharmabot: a pediatric generic medicine consultant chatbot. *J Autom Cont Eng* 2015 Apr;3(2):137-140. [doi: [10.12720/joace.3.2.137-140](https://doi.org/10.12720/joace.3.2.137-140)]
33. Kowatsch T, Volland D, Shih I, Rügger D, Künzler F, Barata F, et al. Design and evaluation of a mobile chat app for the open source behavioral health intervention platform MobileCoach. In: *Designing the Digital Transformation*. Cham: Springer; 2017.
34. L'Allemand D, Shih C, Heldt K, Buchter D, Brogle B, Ruegger D. Design and interim evaluation of a smartphone app for overweight adolescents using a behavioural health intervention platform. *Obes Rev* 2018:102 [FREE Full text]
35. 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 2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. 2018 Presented at: 2018 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)]
36. Ta V, Griffith C, Boatfield C, Wang X, Civitello M, Bader H, et al. User experiences of social support from companion chatbots in everyday contexts: thematic analysis. *J Med Internet Res* 2020 Mar 06;22(3):e16235 [FREE Full text] [doi: [10.2196/16235](https://doi.org/10.2196/16235)] [Medline: [32141837](https://pubmed.ncbi.nlm.nih.gov/32141837/)]
37. 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/)]

Abbreviations

AI: artificial intelligence

CORQ: Consolidated Criteria for Reporting Qualitative Research

mHealth: mobile health

Edited by G Eysenbach; submitted 14.05.21; peer-reviewed by R Batten, MA Bahrami, J Watterson; comments to author 18.08.21; revised version received 09.09.21; accepted 16.09.21; published 11.11.21.

Please cite as:

Dhinakaran DA, Sathish T, Kowatsch T, Griva K, Best JD, Tudor Car L

Public Perceptions of Diabetes, Healthy Living, and Conversational Agents in Singapore: Needs Assessment

JMIR Form Res 2021;5(11):e30435

URL: <https://formative.jmir.org/2021/11/e30435>

doi: [10.2196/30435](https://doi.org/10.2196/30435)

PMID: [34762053](https://pubmed.ncbi.nlm.nih.gov/34762053/)

©Dhakshenya Ardhithy Dhinakaran, Thirunavukkarasu Sathish, Tobias Kowatsch, Konstadina Griva, James Donovan Best, Lorainne Tudor Car. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.11.2021. 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

Seeking Mental Health Support Among College Students in Video-Based Social Media: Content and Statistical Analysis of YouTube Videos

Bogeum Choi¹, MSc; Heejun Kim², PhD; Jina Huh-Yoo³, PhD

¹School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

²Department of Information Science, University of North Texas, Denton, TX, United States

³College of Computing and Informatics, Drexel University, Philadelphia, PA, United States

Corresponding Author:

Heejun Kim, PhD

Department of Information Science

University of North Texas

3940 North Elm, Suite E295G

Denton, TX, 76203

United States

Phone: 1 940 369 8285

Email: heejun.kim@unt.edu

Abstract

Background: Mental health is a highly stigmatized disease, especially for young people. Due to its free, accessible format, college students increasingly use video-based social media for many aspects of information needs, including how-to tips, career, or health-related needs. The accessibility of video-based social media brings potential in supporting stigmatized contexts, such as college students' mental health. Understanding which kinds of videos about college students' mental health have increased viewer engagement will help build a foundation for exploring this potential. Little research has been done to identify video types systematically, how they have changed over time, and their associations on viewer engagement both short term and long term.

Objective: This study aims to identify strategies for using video-based social media to combat stigmatized diseases, such as mental health, among college students. We identify who, with what perspective, purpose, and content, makes up the videos available on social media (ie, YouTube) about college students' mental health and how these factors associate with viewer engagement. We then identify effective strategies for designing video-based social media content for supporting college students' mental health.

Methods: We performed inductive content analysis to identify different types of YouTube videos concerning college students' mental health (N=452) according to video attributes, including poster, perspective, and purpose. Time analysis showed how video types have changed over time. Fisher's exact test was used to examine the relationships between video attributes. The Mann-Whitney *U* test was used to test the association between video types and viewer engagement. Lastly, we investigated the difference in viewer engagement across time between two major types of videos (ie, individuals' storytelling and organization's informational videos).

Results: Time trend analysis showed a notable increase in the number of (1) videos by individuals, (2) videos that represent students' perspectives, and (3) videos that share stories and experiential knowledge over the recent years. Fisher's exact test found all video attributes (ie, poster, perspective, and purpose) are significantly correlated with each other. In addition, the Mann-Whitney *U* test found that poster (individual vs organization) and purpose (storytelling vs sharing information) type has a significant association with viewer engagement ($P<.001$). Lastly, individuals' storytelling videos had a greater engagement in the short term and the long term.

Conclusions: The study shows that YouTube videos on college students' mental health can be well differentiated by the types of posters and the purpose of the videos. Taken together, the videos where individuals share their personal stories, as well as experiential knowledge (ie, tips and advice), engaged more viewers in both the short term and long term. Individuals' videos on YouTube showed the potential to support college students' mental health in unique ways, such as providing social support, validating experience, and sharing the positive experience of help-seeking.

KEYWORDS

mental health; college student; social media; YouTube; help-seeking; experiential knowledge; video types; content analysis; time distribution analysis; depression; anxiety; student; knowledge; stigma; strategy; engagement

Introduction

Preface

Postsecondary education represents a peak onset period for mental disorders [1]. Studies showed mental health disorders affect 12%-46% of all university students at any given year [2-4]. However, due to stigma, embarrassment, and preference for self-reliance, only 1 in 5 receives minimally adequate treatment [3]. On the other hand, young adults look for social support and encouragement from others as aids for the help-seeking process [5].

Studies showed college student communities rely heavily on YouTube for both academic and entertainment purposes [6,7]. YouTube, the most popular video-sharing website and the second most popular social media website after Facebook, had 2 billion users as of 2020. A new video is uploaded every minute, and, on average, a user spends at least 15 minutes a day on the site [8,9]. As a result, researchers have been increasingly examining YouTube as a source of health information, including descriptions of various health conditions such as rheumatoid arthritis, weight loss, infertility, anxiety, acute myocardial infarction, and attention deficit disorder [10]. In these studies, the analysis focused on the validity of shared content as factual information about the diseases, prognosis, and symptoms. While misinformation is a significant challenge in social media, one of the benefits of social media is peer support and the personal experiences of others that may be difficult to access otherwise, mainly due to the stigma around sharing about one's mental illness [11,12].

Studies showed videos containing personal narratives and experiential knowledge are associated with higher viewer engagement and are preferred sources for seeking mental health information [13]. However, one thing that remains unclear about this type of video is their growth rate pattern [14], which is a critical measure for engagement and can be predictive of popularity. No studies of YouTube videos on health have examined growth patterns and what it means to generate constructive and engaging videos that serve as long-lived resources for the target audience. Despite the imminent need, studies have underexplored how video-based social media, such as YouTube, provide a supportive environment for stigmatized diseases such as mental health among college students.

This work investigates how social media provides an environment in which individuals who are susceptible to stigma can benefit from others' personal narratives and experiential information. To address this goal, we first conducted a systematic search of YouTube videos. We then performed a qualitative content analysis to identify key categories for video attributes such as poster, perspective, and purpose. Then we looked at a distribution of the video attributes over time. Finally, using the categories developed from the content analysis, we

performed statistical analyses (ie, Fisher's exact test and Mann-Whitney *U* test) to investigate (1) the relationships among the video attributes and (2) the association between video attributes and viewer engagement measures such as views, likes, dislikes, and comments.

The findings inform our understanding of the videos published about college students' mental health, such as what types of videos are published on college students' mental health, which has become more common over time, and how viewers are engaging them. This finding is essential for informing future work and provides insights for improving some videos for greater engagement.

Background

College students' mental health has posed a significant concern in higher education. A WHO (World Health Organization) survey in 19 colleges across 8 countries showed nearly 35% (4894/13984) of college student interviewees met DSM-IV criteria for at least one mental disorder in lifetime and 31% (4335/13984) in the last 12 months, where anxiety and mood disorders were the most prevalent [15]. Other studies have reported similar or even higher prevalence rates of mental disorders among college students [1,16-18]. However, mental health service utilization among this population can be low. A survey study showed, among students with mental disorders, only 36% received any treatment in the past year [19]. The barriers to receiving the treatment include stigma, embarrassment, and preference for self-reliance [5]. On the other hand, students are more open to receiving help when the help is provided through social support and encouragement from peers [20,21]. Due to these reasons, college students are more likely to search for informal alternatives, such as going online [22,23].

Peer support exchanged through social media has received much attention in the research community. By sharing personal experiences and providing emotional support among those with a similar background, researchers examined how internet-based peer support through social media (eg, community forum and YouTube) can positively impact behavior change and health outcomes [24]. Peer engagement and social support can facilitate college students' help-seeking for mental health issues. Peer support on social media, combined with college students' heavy use of YouTube [6,7], presents the potential of YouTube videos to encourage college students to acknowledge their mental health issues and seek and receive help. However, there is a lack of investigation on video content, genre, and viewer engagement, especially concerning engagement over time. Therefore, it is critical to investigate how YouTube videos can support this population (ie, what kinds of videos are available related to college students' mental health and how the viewers have engaged these videos over time). Understanding the content and the types of videos that effectively engage people will help

identify and evaluate ways to bring support mechanisms through nonclinical means of communication such as social media.

We conducted a content analysis and statistical analysis of YouTube videos related to college students' mental health and their engagement metrics to address this gap. In addition, we discuss implications for how organizations, groups, and individuals can use video-based social media platforms to support stigmatized conditions that can benefit from peer support.

Research Question

RQ1: What types of videos are available on YouTube regarding college students' mental health?

We investigate the video types available on YouTube regarding college students' mental health by identifying the following attributes of a video: poster, perspective, and purpose. We also examine how the distribution of the identified categories for each video attribute changes over time and if there is a systematic relationship between the video attributes.

RQ2: Do video types have associations with viewer engagement?

We investigate whether the types of video, as defined by video attributes, have any associations with viewer engagement measured by the number of views, likes, dislikes, and comments. We further examine if a video type (ie, individual's storytelling video vs organization's informational video) has an association with long-term viewer engagement.

Textbox 1. Categories and search keywords.

Illness: Depression, anxiety, stress, mental health

Population: College student, university student, campus

Context: How to, life

Examples of search query: Depression college student how to, anxiety university student how to, mental health campus life, etc

One word from each category was used at a time to construct a query.

The search was conducted between May 28, 2020, and May 30, 2020, which resulted in 1930 videos. After removing duplicates, 860 unique videos were obtained. Next, 2 authors screened the video's title and description based on the exclusion criteria (ie, the link is no longer active or not in English) and the inclusion criteria (ie, the video is about college students and mental health). This screening process yielded 452 videos. The 408 videos eliminated during the screening include students preparing for college rather than students enrolled in college or those about general health, not mental health.

Data Analysis

Content Analysis

To address RQ1, we conducted an iterative, inductive content analysis [26]. We identified and categorized relevant information

Methods

Data Collection

We conducted a keyword search of YouTube videos concerning the mental health of college students. We formulated search keywords based on three topical categories: illness, population, and context. For the illness-related keyword, we selected 4 top symptoms of college students receiving mental health services from the 2019 Center for Collegiate Mental Health Annual Report, including depression, anxiety, social anxiety, and academic distress. Based on Consumer Health Vocabulary, we changed these into more commonly used terms, "depression," "anxiety," and "stress" [25]. We also included an overarching term, "mental health," in this category. For the population-related keyword, we used "college student(s)" and other terms including "university student(s)" and "campus" after reviewing related search terms that YouTube suggested with "college student(s)." We extracted context-related keywords from the initial search results retrieved by illness and population keywords. The goal was to reflect a variety of contexts in which the videos can be sought and used by college students. We then combined the terms from each category (illness, population, and context) and used it as a search query (see [Textbox 1](#)).

To alleviate biases from Google's personalization algorithms, we developed a Python (version 3.8.5; Python Software Foundation) script that drives an Incognito Chrome browser and enables anonymous searches from YouTube. The script ran all the possible combinations of keywords as a query and parsed video IDs from retrieved videos. These IDs were then passed over to the YouTube API to collect metadata (eg, video description, publication date, views, likes, dislikes, and comments) associated with each video.

on 3 specific video attributes: (1) poster, (2) perspective, and (3) purpose. The analysis was conducted using the title, written description, and the content of the video. The analysis was led by 1 coder (first author), while the whole team of 4 researchers iteratively discussed the codes and reached a consensus on the most salient code names and descriptions from the data. The coder then revisited the originally coded data and recoded those data using the revised coding scheme developed as a team. We iterated this process until the team agreed upon a final set of code names and definitions. Then we identified categories of related themes across the codes. These categories will be presented as findings for each video attribute.

Statistical Analysis

First, we applied time distribution analysis to demonstrate how the distribution of categories of each video attribute changes over time. We plotted the occurrences of the categories of video attributes by year. Additionally, we performed Fisher's exact test to investigate the relationship between video attributes. We used Fisher's exact test instead of the chi-square test because our data had categories with a low-frequency value, and using the chi-square test with low cell counts (ie, less than 5) could lead to increased type I errors.

To address RQ2, we first performed the Mann-Whitney *U* test to examine the association between video attributes and viewer engagement scores (ie, views, comments, likes, and dislikes). The Mann-Whitney *U* test, also known as Wilcoxon rank-sum test, can test differences between 2 groups when the data is not normally distributed. Specifically, we examined the following attributes: poster (individual vs organization), perspective (student vs domain expert), purpose (storytelling vs sharing information), and poster combined with the purpose (individual + storytelling vs organization + sharing information). We compared these specific categories because those were the main, distinct categories that make up the videos on college students' mental health by the content analysis. The viewer engagement scores were normalized to control the confounding effects of exposure time. In other words, the scores (eg, the number of views) were divided by each video's exposure time. We used R (version 3.6.2; R Core Team) for all statistical analyses.

Lastly, we demonstrated the association between video attributes (specified above) and long-term viewer engagement. This analysis chose the number of comments from the 4 engagement scores as a dependent variable because we believed it to be a more meaningful measure for viewers' engagement with a video. Prior work also suggests that commenting reflects a higher engagement level than viewing or liking as it involves expressing one's opinions or feelings in public [27]. Since we had uneven distribution between the groups to be compared (eg, more videos by individuals than organizations), we randomly selected the same number of videos for each group from those published between 2013 and 2019. Then we compared the number of comments in the years following the publication year ($n+0$, $n+1$, $n+2$... where n is the publication year).

Results

Below, we discuss the findings in response to the two research questions:

1. RQ1: What types of videos are available on YouTube regarding college students' mental health?
2. RQ2: Do video types have an association with viewer engagement?

Finding 1. Video Types on YouTube About College Students' Mental Health

Overview

We first present categories identified for the following video attributes: poster, perspective, and purpose of a video. A relative frequency (the portion of times a category occurred in our data)

is presented in parentheses. We then show how the distribution of video types has changed over time based on video attributes. Lastly, we present how the video attributes are correlated with each other.

Video Attributes

Poster

We identified 5 categories for poster types: (1) individual (55%), (2) media (19%), (3) university (18%), (4) health-related organization (6%), and (5) other organization (2%). Individual was coded when a video was created and published by a single individual or a small group of individuals such as a student group. Many of these individuals were college students who either have experienced or are currently experiencing mental health issues. Media included news outlets, television shows, and professional content production groups (eg, TED). University referred to the university itself and university-affiliated organizations that provide mental health-related services to students (eg, student counseling center).

Perspective

We identified 5 mutually inclusive categories for the purpose of the videos. Student's perspective (74%) was the most often represented in the videos, followed by domain expert (32%), others (17%), and amateur coach (7%). Domain expert included mental health experts or those who have expertise in advising students. Amateur coach refers to people who play a role of an advisor based on their life experience while not having credentials in a traditional sense. For example, a public speaker or a podcast host who offers advice on mental health issues was coded as amateur coach.

Purpose

We identified 5 mutually inclusive categories for the purpose of the videos. Sharing experiential knowledge was the most frequently identified (58%), followed by sharing information (42%), sharing stories (39%), promoting help-seeking (23%), and sharing awareness (16%). Sharing experiential knowledge was coded when a person in a video shares what they learned from their first-hand experience of mental health. The most common theme under sharing experiential knowledge was sharing tips and advice. Those tips and advice include "strategies to cope with stress and anxiety" (V304), "how to beat depression in college" (V26), "how to deal with stress, anxiety, and depression" (V31), etc. Experiential knowledge is distinct from factual knowledge in 3 significant ways. First, it is drawn from people's lived experiences, thus more relatable to someone undergoing similar issues. Second, it is often shared to help, support, or guide others. Third, it has a practical value in that viewers can incorporate such tips and advice into the everyday practice of their lives. For example, a person in V312 stated,

...so, I'm going to share with you guys what worked for me, and hopefully, you guys might be able to take some sort of advice from it. Maybe even adopt it and change it into what works for your life.

Sharing information included videos that deliver factual information about mental health (eg, symptoms, causes and treatments, statistics) within and outside an educational setting.

Sharing stories included videos that share personal mental health stories in a narrative style. Some common themes under sharing stories included offering emotional support and sharing an ordinary day of a college student with a mental health issue. Offering emotional support referred to cases where a person in a video showed emotional support for their viewers while disclosing their personal issues. For example, the poster of V56 said,

I opened up to you guys a bit more than I ever have before; I was hesitant to keep that part in the video but decided to push myself out of my comfort zone because hopefully, it helps someone know they're not alone.

In addition, we saw college students share their ordinary days to demonstrate what it is like to be a college student with a mental health issue. For example, the poster of V303 had a series called “uni vlog,” where he showed how he reacted to and coped with a stressful situation in a vlog format. Another poster (V81) discussed her motivation for vlogging as

You can gain a perspective as to what it is for one kid who goes through this and that [...] My main goal is to create a vast database of experiences for people to see so they can learn before they have to make the same mistake.

Promoting help-seeking included videos that conveyed help-seeking messages. Help-seeking included not only getting professional help but also talking to friends or family about mental health problems. For example, a poster of V235 said,

Don't be afraid to ask for help. We need to stop feeling like we are super-human. We are really not.

Posters also encouraged others to seek help by sharing their experience of getting professional help during their mental health journey. For instance, the poster of V59 said,

I refused to go to therapy for many years...but looking back...No. Go to therapy. Talk about your emotions. You have to get it out. You are going to have to process those emotions eventually, so just start trying. Go to therapy as much as possible. I promise, regardless of your personal type, it helps.

Similarly, the poster of V48 said,

If you have depression, or you suspect that you have depression, I'd encourage you to try to reach out to a therapist. Maybe you need to just talk to someone. Believe me, I thought that I would never reach out to get professional help. I thought that was completely out of the question. I'm just telling you, if I can do it, you can do it. Trust me.

Video Types Over Time

Based on the identified categories for each video attribute, we looked at how the types of videos posted on YouTube about mental health have changed over time. First, there was a significant increase in videos posted by an individual compared to an organization (Figure 1). Second, the number of videos representing students’ perspectives has increased over the years compared to other perspectives, such as domain experts and amateur coaches (Figure 2). Third, the number of storytelling videos has shown the most significant increase among all-purpose types in the past few years (Figure 3). While educational videos were the most prevalent before 2015, videos sharing personal stories and experiential knowledge have surpassed them since 2016.

Figure 1. Type of poster by year.

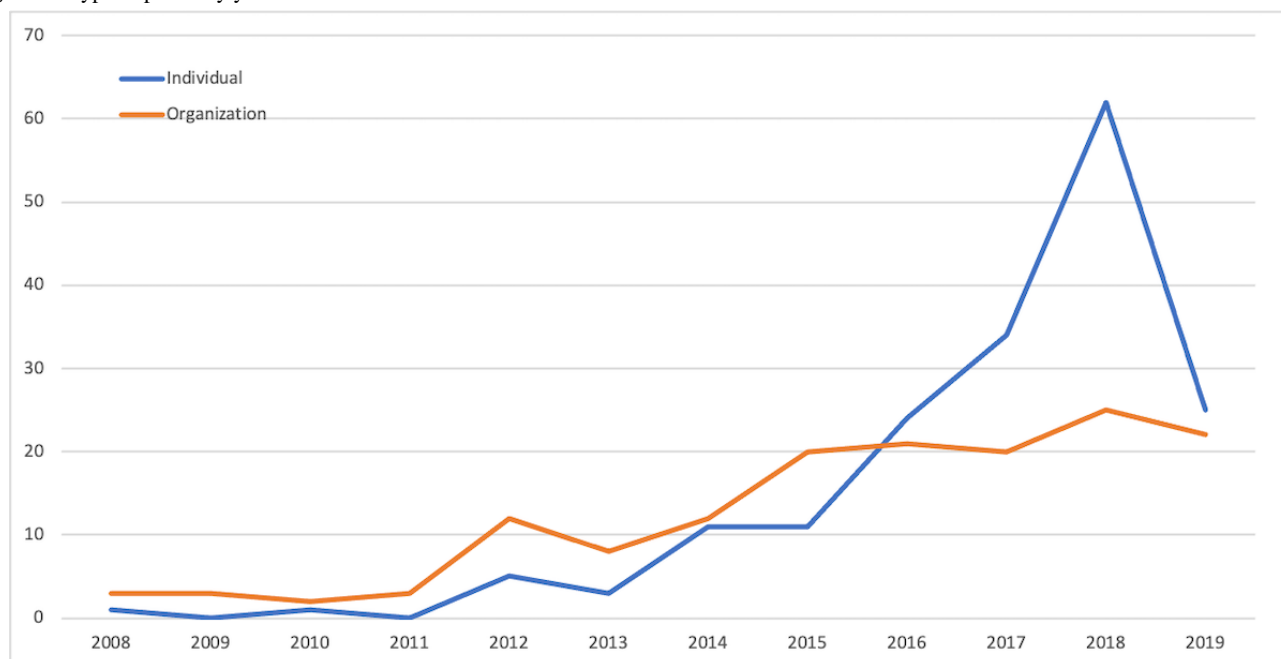


Figure 2. Type of perspective by year.

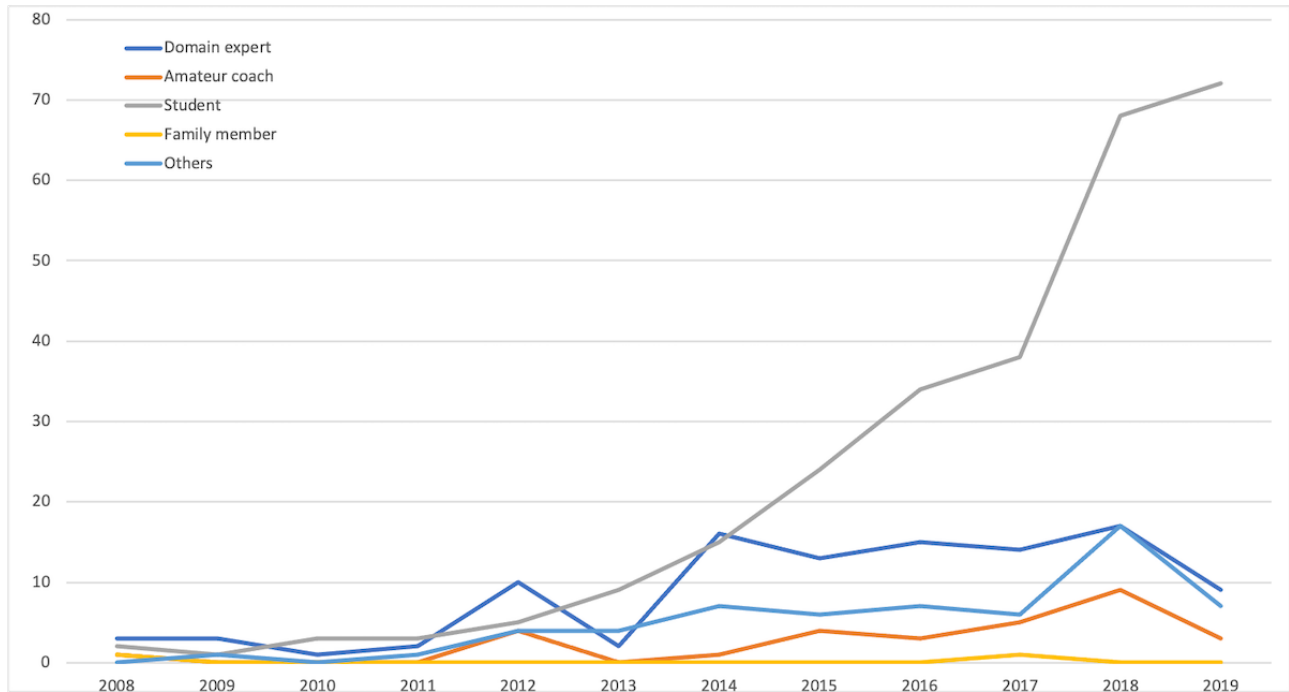
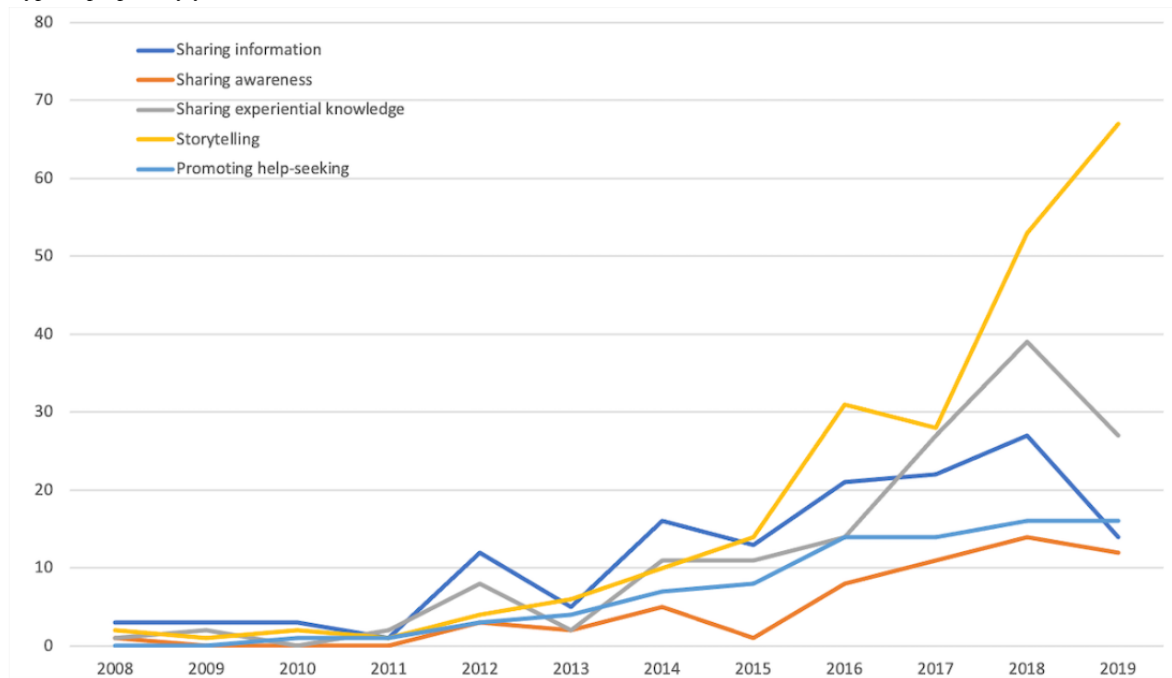


Figure 3. Type of purpose by year.



Relationships Between Video Attributes

We further investigated the relationships between video attributes to see if specific attributes tend to co-occur more. According to Fisher’s exact test, all combinations of our video attributes showed a significant relationship with each other. First, poster type significantly correlated with the perspective type ($P < .001$). As shown in Table 1, individuals’ videos were more likely to include students’ perspectives in their videos, while organizations’ videos were likely to include both student and domain expert perspectives. Poster type and purpose type were also correlated ($P = .03$). As shown in Table 2, individuals’

videos primarily focused on storytelling and sharing experiential knowledge, while organizations’ videos focused on sharing information. Lastly, perspective type significantly correlated with the purpose type ($P < .001$). As shown in Table 3, videos representing students’ perspectives share stories and experiential knowledge, while videos with domain experts’ perspectives share information. Each cell in Tables 1- 3 presents the number of co-occurring attributes and the percentage in parenthesis. As mentioned above, the categories for perspective and purpose were not mutually exclusive, which means a video could be coded as more than one category.

Table 1. Number and percentage of poster versus perspective.

Poster	Perspective				
	Student (n=293)	Domain expert (n=116)	Amateur coach (n=31)	Family member (n=2)	Others (n=65)
Individual (n=251), n (%)	189 (47.6)	31 (7.8)	16 (4.0)	0 (0)	15 (3.8)
Organization (n=258), n (%)					
Media (n=123)	47 (11.8)	32 (8.1)	8 (2.0)	2 (0.5)	34 (8.6)
University (n=94)	44 (11.1)	34 (8.6)	4 (1.0)	0 (0)	12 (3.0)
Health-related org (n=27)	8 (2.0)	16 (4.0)	2 (0.5)	0 (0)	1 (0.3)
Other org (n=14)	5 (1.3)	3 (0.8)	1 (0.3)	0 (0)	3 (0.8)

Table 2. Number and percentage of poster versus purpose.

Poster	Purpose				
	Share experiential knowledge (n=158)	Share information (n=155)	Share stories (n=237)	Promote help-seeking (n=93)	Share awareness (n=64)
Individual (n=393), n (%)	107 (27.0)	52 (13.1)	165 (41.6)	39 (9.8)	30 (7.6)
Organization (n=314), n (%)					
Media (n=138)	19 (4.8)	52 (13.1)	36 (9.1)	15 (3.8)	16 (4.0)
University (n=126)	20 (5.0)	31 (7.8)	29 (7.3)	33 (8.3)	13 (3.3)
Health-related org (n=38)	8 (2.0)	17 (4.3)	5 (1.3)	4 (1.0)	4 (1.0)
Other org (n=12)	4 (1)	3 (0.8)	2 (0.5)	2 (0.5)	1 (0.3)

Table 3. Number and percentage of perspective versus purpose.

Perspective	Purpose, n (%)				
	Share experiential knowledge (n=190)	Share information (n=248)	Share stories (n=283)	Promote help-seeking (n=122)	Share awareness (n=96)
Student (n=539)	115 (12.2)	77 (8.2)	227 (24.2)	68 (7.2)	52 (5.5)
Domain expert (n=232)	39 (4.2)	97 (10.3)	32 (3.4)	35 (3.7)	29 (3.1)
Amateur coach (n=51)	23 (2.4)	13 (1.4)	7 (0.7)	7 (0.7)	1 (0.1)
Family member (n=6)	0 (0)	2 (0.2)	2 (0.2)	0 (0)	2 (0.2)
Other (n=111)	13 (1.4)	59 (6.3)	15 (1.6)	12 (1.3)	12 (1.3)

Finding 2. The Association of Video Type With Viewer Engagement

This section presents the association between video types and viewer engagement measured by the number of views, likes, dislikes, and comments. Content analysis informed us that the 2 most distinct types of YouTube videos regarding college students' mental health could be characterized by the poster and purpose attributes. Thus, we used poster, purpose, and a combination of those two attributes to define video types to compare.

The Association Between Video Type and Viewer Engagement

According to the Mann-Whitney *U* test, type of poster, purpose, and poster combined with purpose had significant differences on all measures of viewership engagement (Table 4). Specifically, individuals' videos (vs organizations' videos) and storytelling videos (vs informational videos) demonstrated significantly higher viewer engagement than their counterparts. Not surprisingly, individuals' storytelling videos also demonstrated higher viewer engagement than organizations' information videos.

Table 4. Viewer engagement by video types.

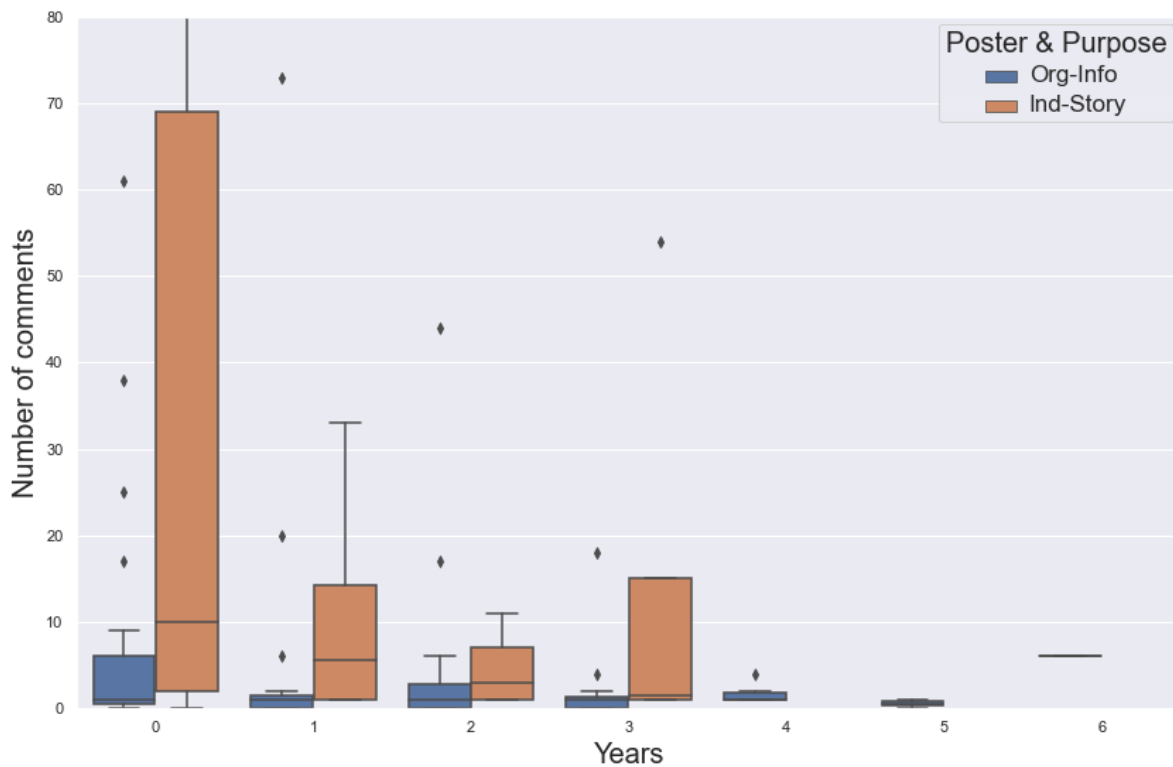
Compared video types	Engagement scores			
	View count	Like count	Dislike count	Comment count
Poster type (individual vs organization)				
Individual, median	3.120	0.096	0.001	0.020
Organization, median	1.011	0.008	0.000	0.000
Mann-Whitney <i>U</i>	14303.5	10374.0	14214.5	9449.5
<i>P</i> value	<.001	<.001	<.001	<.001
Purpose type (sharing stories vs sharing information)				
Storytelling, median	2.790	0.061	0.001	0.011
Sharing information, median	0.681	0.005	0.000	0.000
Mann-Whitney <i>U</i>	12853.0	10207.5	14174.0	10563.0
<i>P</i> value	<.001	<.001	<.001	<.001
Poster + purpose (individual + sharing stories vs organization + sharing information)				
Individual + sharing, median	3.502	0.131	0.002	0.030
Organization + sharing information, median	0.744	0.005	0.000	0.000
Mann-Whitney <i>U</i>	5446.0	3482.0	5943.5	3257.5
<i>P</i> value	<.001	<.001	<.001	<.001

The Association Between Video Type and Viewer Engagement Over Time

Figure 4 shows how the association between video type (ie, individual’s storytelling vs organization’s informational videos)

and viewer engagement varies over time. Individuals’ storytelling videos had a significantly higher engagement at the time of posting and continued to attract more engagements than organizations’ informational videos after some time since posting.

Figure 4. Number of comments by poster-purpose type (individual-sharing stories versus organization-sharing information).



Discussion

Principal Findings

This study sought to understand the videos available on YouTube regarding college students' mental health (RQ1) and the relationships between video types and viewer engagement scores (RQ2).

We found an increasing number of individuals posting videos on YouTube to share their experiential knowledge and personal stories related to mental health. The trend has gotten more robust over time. Individuals' storytelling videos were associated with greater engagement (ie, views, likes, dislikes, and comments) than organizations' informational videos. Moreover, individuals' storytelling videos seem to continue getting engagement (in terms of the number of comments) sometime after the posting. This finding might indicate that users seeking mental health-related videos about college students on YouTube may want to see and hear from those in a similar situation talking about their first-hand experience. Indeed, previous studies found that people use social media platforms for peer support [28,29]. In our data, individuals' storytelling videos seem to meet such needs among college students, especially considering that many of the storytelling videos are intended to offer emotional support and share everyday life from the first-person perspective. However, we need a careful examination to know the causal relationship between video types and viewer engagement scores.

Personal stories shared in individuals' videos can have other specific benefits. Many of these videos included individuals' recovery narratives. Previous research noted positive outcomes of patients receiving recovery narratives, including connectedness, hope, stigma reduction [30], and validation of experience [31]. Lack of a benchmark can hinder people from recognizing their symptoms, resulting in a barrier to help-seeking [32]. In this regard, the first-hand experience shared in individuals' videos on YouTube can serve as a reference point while providing validation to college student viewers with mental health issues. This approach can be especially beneficial to individuals whose cultural background may be more prone to the stigma around decreased mental health than others, such as African-American [33] or Asian populations [34]. This benefit will be sustainable if YouTube, as a platform, provides culturally sensitive, accessible information and experience sharing in those populations.

We also found that individuals' videos frequently share experiential knowledge, which referred to experience-based tips and advice. This content includes activities and mindsets that one can incorporate into their daily lives to improve mental health or better cope with a specific mental health issue (eg, social anxiety). Previous research suggests that people choose to use social media platforms instead of search engines to seek other people's recommendations and advice on how to manage mental health in particular [35,36] and learn from it, especially among young people [37]. The growing prevalence of videos that share experiential knowledge suggests a potential benefit of giving college students viewers access to other people's lived experiences.

It is worth noting that the portion of individuals' videos that promote help-seeking (39/397, 9.8%) was comparable with that of organizations' videos (54/397, 13.5%) in this study. Previous research found that patients' perceived treatment outcome is prone to social influence from other patients' shared experiences [38]. Moreover, our study shows that individuals' videos get more engagement than organizations'. This finding suggests that individuals' videos promoting help-seeking while sharing their mental health journey could influence viewers to have a positive attitude towards help-seeking, including formal mental health clinics and services.

Our findings suggest that organizations (eg, universities, health-related organizations) may benefit from incorporating the communication styles used in individuals' videos when reaching college students about mental health on a large scale. For instance, universities can promote their mental health service by including individual students' anecdotes about their experience using the service to improve their mental health. Some organizations' videos included clips of students being interviewed in our data, but many of them were either too short or formal. Organizations may gain greater engagement and increase help-seeking among college student viewers by incorporating more informal personal narratives in delivering mental health-related content.

Strengths and Limitations

Our study has several strengths and limitations. We provided multiple pieces of evidence that allude to the growing popularity and engagement of individuals' storytelling videos, including the changing prevalence of videos with different attributes and the association between video types and viewer engagement. As far as we know, no previous research has investigated viewer engagement with videos on the topics of college students' mental health over time. We included this analysis to see if the trend of individuals' storytelling videos getting more engagement holds over time. We then subsequently discussed how those individuals' videos could support college students with mental health issues in more detail based on the qualitative analysis conducted to identify categories of video attributes. We used quantitative measures such as the number of views, likes, dislikes, and comments. Although these measures available from YouTube allow us to gauge the popularity and accessibility of videos, they do not inform if and how videos impact viewers' mental health. To investigate the actual impact of videos with specific attributes, in our future work, we plan to expand our data collection to include college students' responses to different types of videos.

Conclusions

Our study defined different mental health-related videos intended for college students based on video attributes (eg, poster and purpose) and examined engagement with those different video types. Our findings show individuals' videos that share stories and experiential knowledge have become more popular over time and tend to gain more engagement from viewers for both the short term and long term. Such videos may attract and sustain engagement due to their benefits to college students with mental health issues, including peer support, validating experience, learning from others' experiences, and

encouraging help-seeking. While the quality of health information on social media platforms is debatable, it is essential to consider that there are different types of mental health needs of college students that can be more effectively addressed by individually-generated content on social media such as

YouTube. When making interventions targeting help-seeking for mental health issues among college students, incorporating the communication styles used by individuals on social media such as storytelling will be strategic.

Acknowledgments

We acknowledge Chris Yang, Sunyoung Park, and Eugenia Chan for providing resources and ideas for enabling this work. Furthermore, we thank William Sutherland-Keller for manuscript proofing.

Conflicts of Interest

None declared.

References

1. Ibrahim AK, Kelly SJ, Adams CE, Glazebrook C. A systematic review of studies of depression prevalence in university students. *J Psychiatr Res* 2013 Mar;47(3):391-400. [doi: [10.1016/j.jpsychires.2012.11.015](https://doi.org/10.1016/j.jpsychires.2012.11.015)] [Medline: [23260171](https://pubmed.ncbi.nlm.nih.gov/23260171/)]
2. Auerbach RP, Alonso J, Axinn WG, Cuijpers P, Ebert DD, Green JG, et al. Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychol Med* 2016 Oct;46(14):2955-2970 [FREE Full text] [doi: [10.1017/S0033291716001665](https://doi.org/10.1017/S0033291716001665)] [Medline: [27484622](https://pubmed.ncbi.nlm.nih.gov/27484622/)]
3. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology* 2018 Oct;127(7):623-638. [doi: [10.1037/abn0000362](https://doi.org/10.1037/abn0000362)] [Medline: [30211576](https://pubmed.ncbi.nlm.nih.gov/30211576/)]
4. Verger P, Guagliardo V, Gilbert F, Rouillon F, Kovess-Masfety V. Psychiatric disorders in students in six French universities: 12-month prevalence, comorbidity, impairment and help-seeking. *Soc Psychiatry Psychiatr Epidemiol* 2010 Feb;45(2):189-199. [doi: [10.1007/s00127-009-0055-z](https://doi.org/10.1007/s00127-009-0055-z)] [Medline: [19381424](https://pubmed.ncbi.nlm.nih.gov/19381424/)]
5. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry* 2010 Dec 30;10:113 [FREE Full text] [doi: [10.1186/1471-244X-10-113](https://doi.org/10.1186/1471-244X-10-113)] [Medline: [21192795](https://pubmed.ncbi.nlm.nih.gov/21192795/)]
6. Fleck BKB, Beckman LM, Sterns JL, Hussey HD. YouTube in the Classroom: Helpful Tips and Student Perceptions. *Journal of Effective Teaching* 2014;14(3):21-37 [FREE Full text]
7. Moghavvemi S, Sulaiman A, Jaafar NI, Kasem N. Social media as a complementary learning tool for teaching and learning: The case of youtube. *The International Journal of Management Education* 2018 Mar;16(1):37-42. [doi: [10.1016/j.ijme.2017.12.001](https://doi.org/10.1016/j.ijme.2017.12.001)]
8. Most popular social networks worldwide as of January 2021, ranked by number of active users Internet. Tankovska H. URL: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/> [accessed 2021-05-27]
9. Most Popular Social Media Platforms in 2021. URL: <https://www.oberlo.com/statistics/most-popular-social-media-platforms> [accessed 2021-05-27]
10. Thapa P, Thapa A, Khadka N, Bhattarai R, Jha S, Khanal A, et al. YouTube lens to attention deficit hyperactivity disorder: a social media analysis. *BMC Res Notes* 2018 Dec 04;11(1):854 [FREE Full text] [doi: [10.1186/s13104-018-3962-9](https://doi.org/10.1186/s13104-018-3962-9)] [Medline: [30514370](https://pubmed.ncbi.nlm.nih.gov/30514370/)]
11. Huh J, Liu LS, Neogi T, Inkpen K, Pratt W. Health Vlogs as Social Support for Chronic Illness Management. *ACM Trans Comput Hum Interact* 2014 Aug;21(4):1-31 [FREE Full text] [doi: [10.1145/2630067](https://doi.org/10.1145/2630067)] [Medline: [26146474](https://pubmed.ncbi.nlm.nih.gov/26146474/)]
12. Seo W, Jung H. Exploring the community of blind or visually impaired people on YouTube. New York, NY, USA: Association for Computing Machinery; 2017 Oct Presented at: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility; 2017; Baltimore, Maryland, USA p. 371-372 URL: <https://dl.acm.org/doi/10.1145/3132525.3134801> [doi: [10.1145/3132525.3134801](https://doi.org/10.1145/3132525.3134801)]
13. Oliphant T. User Engagement with Mental Health Videos on YouTube. *Journal of the Canadian Health Libraries Association* 2013 Dec 09;34(3):153-158. [doi: [10.5596/c13-057](https://doi.org/10.5596/c13-057)]
14. Figueiredo F, Benevenuto F, Almeida JM. The tube over time: Characterizing popularity growth of YouTube videos. New York, NY, USA: Association for Computing Machinery; 2011 Feb Presented at: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining; 2011; Hong Kong, China p. 745-754. [doi: [10.1145/1935826.1935925](https://doi.org/10.1145/1935826.1935925)]
15. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *J Abnorm Psychol* 2018 Oct;127(7):623-638 [FREE Full text] [doi: [10.1037/abn0000362](https://doi.org/10.1037/abn0000362)] [Medline: [30211576](https://pubmed.ncbi.nlm.nih.gov/30211576/)]
16. Blanco C, Okuda M, Wright C, Hasin DS, Grant BF, Liu SM, et al. Mental health of college students and their non-college-attending peers: results from the National Epidemiologic Study on Alcohol and Related Conditions. *Arch Gen Psychiatry* 2008 Dec;65(12):1429-1437 [FREE Full text] [doi: [10.1001/archpsyc.65.12.1429](https://doi.org/10.1001/archpsyc.65.12.1429)] [Medline: [19047530](https://pubmed.ncbi.nlm.nih.gov/19047530/)]

17. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. *Acad Med* 2006 Apr;81(4):354-373. [doi: [10.1097/00001888-200604000-00009](https://doi.org/10.1097/00001888-200604000-00009)] [Medline: [16565188](https://pubmed.ncbi.nlm.nih.gov/16565188/)]
18. Zivin K, Eisenberg D, Gollust SE, Golberstein E. Persistence of mental health problems and needs in a college student population. *J Affect Disord* 2009 Oct;117(3):180-185. [doi: [10.1016/j.jad.2009.01.001](https://doi.org/10.1016/j.jad.2009.01.001)] [Medline: [19178949](https://pubmed.ncbi.nlm.nih.gov/19178949/)]
19. Eisenberg D, Hunt J, Speer N, Zivin K. Mental health service utilization among college students in the United States. *J Nerv Ment Dis* 2011 May;199(5):301-308. [doi: [10.1097/NMD.0b013e3182175123](https://doi.org/10.1097/NMD.0b013e3182175123)] [Medline: [21543948](https://pubmed.ncbi.nlm.nih.gov/21543948/)]
20. Dennis JM, Phinney JS, Chuateco LI. The Role of Motivation, Parental Support, and Peer Support in the Academic Success of Ethnic Minority First-Generation College Students. *Journal of College Student Development* 2005;46(3):223-236. [doi: [10.1353/csd.2005.0023](https://doi.org/10.1353/csd.2005.0023)]
21. Kuh G, Kinzie J, Schuh JH, Whitt EJ. *Student Success in College: Creating Conditions That Matter*. Hoboken, NJ: Jossey-Bass; Jun 2010:416.
22. Sheffield JK, Fiorenza E, Sofronoff K. Adolescents' Willingness to Seek Psychological Help: Promoting and Preventing Factors. *Journal of Youth and Adolescence* 2004 Dec;33(6):495-507. [doi: [10.1023/B:JOYO.0000048064.31128.c6](https://doi.org/10.1023/B:JOYO.0000048064.31128.c6)]
23. Gowen LK. Online Mental Health Information Seeking in Young Adults with Mental Health Challenges. *Journal of Technology in Human Services* 2013 Apr;31(2):97-111. [doi: [10.1080/15228835.2013.765533](https://doi.org/10.1080/15228835.2013.765533)]
24. Mo PKH, Coulson NS. Living with HIV/AIDS and use of online support groups. *J Health Psychol* 2010 Apr;15(3):339-350. [doi: [10.1177/1359105309348808](https://doi.org/10.1177/1359105309348808)] [Medline: [20348355](https://pubmed.ncbi.nlm.nih.gov/20348355/)]
25. Zeng QT, Tse T. Exploring and Developing Consumer Health Vocabularies. *Journal of the American Medical Informatics Association* 2006 Jan 01;13(1):24-29. [doi: [10.1197/jamia.m1761](https://doi.org/10.1197/jamia.m1761)]
26. Strauss A, Corbin JM. *Basics of qualitative research: Grounded theory procedures and techniques*. - Pscynet Internet. Newbury Park, CA: Sage Publications, Inc; 1990.
27. Aldous KK, An J, Jansen BJ. View, Like, Comment, Post: Analyzing User Engagement by Topic at 4 Levels across 5 Social Media Platforms for 53 News Organizations. 2019 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; June 11-14, 2019; Munich, Germany p. 47-57 URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/3208>
28. Benetoli A, Chen TF, Aslani P. Consumer Health-Related Activities on Social Media: Exploratory Study. *J Med Internet Res* 2017 Oct 13;19(10):e352 [FREE Full text] [doi: [10.2196/jmir.7656](https://doi.org/10.2196/jmir.7656)] [Medline: [29030326](https://pubmed.ncbi.nlm.nih.gov/29030326/)]
29. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. *Epidemiol Psychiatr Sci* 2016 Apr;25(2):113-122 [FREE Full text] [doi: [10.1017/S2045796015001067](https://doi.org/10.1017/S2045796015001067)] [Medline: [26744309](https://pubmed.ncbi.nlm.nih.gov/26744309/)]
30. Rennick-Egglestone S, Morgan K, Llewellyn-Beardsley J, Ramsay A, McGranahan R, Gillard S, et al. Mental Health Recovery Narratives and Their Impact on Recipients: Systematic Review and Narrative Synthesis. *Can J Psychiatry* 2019 Oct;64(10):669-679 [FREE Full text] [doi: [10.1177/0706743719846108](https://doi.org/10.1177/0706743719846108)] [Medline: [31046432](https://pubmed.ncbi.nlm.nih.gov/31046432/)]
31. Pretorius C, Chambers D, Cowan B, Coyle D. Young People Seeking Help Online for Mental Health: Cross-Sectional Survey Study. *JMIR Ment Health* 2019 Aug 26;6(8):e13524 [FREE Full text] [doi: [10.2196/13524](https://doi.org/10.2196/13524)] [Medline: [31452519](https://pubmed.ncbi.nlm.nih.gov/31452519/)]
32. Biddle L, Donovan J, Sharp D, Gunnell D. Explaining non-help-seeking amongst young adults with mental distress: a dynamic interpretive model of illness behaviour. *Sociol Health Illn* 2007 Nov;29(7):983-1002 [FREE Full text] [doi: [10.1111/j.1467-9566.2007.01030.x](https://doi.org/10.1111/j.1467-9566.2007.01030.x)] [Medline: [18092979](https://pubmed.ncbi.nlm.nih.gov/18092979/)]
33. Planey AM, Smith SMN, Moore S, Walker TS. Barriers and facilitators to mental health help-seeking among African American youth and their families: A systematic review study. *Children and Youth Services Review* 2019 Jun;101:190-200. [doi: [10.1016/j.childyouth.2019.04.001](https://doi.org/10.1016/j.childyouth.2019.04.001)]
34. Zhang Z, Sun K, Jatchavala C, Koh J, Chia Y, Bose J, et al. Overview of Stigma against Psychiatric Illnesses and Advancements of Anti-Stigma Activities in Six Asian Societies. *Int J Environ Res Public Health* 2019 Dec 31;17(1):280 [FREE Full text] [doi: [10.3390/ijerph17010280](https://doi.org/10.3390/ijerph17010280)] [Medline: [31906068](https://pubmed.ncbi.nlm.nih.gov/31906068/)]
35. Aref-Adib G, O'Hanlon P, Fullarton K, Morant N, Sommerlad A, Johnson S, et al. A qualitative study of online mental health information seeking behaviour by those with psychosis. *BMC Psychiatry* 2016 Jul 11;16:232 [FREE Full text] [doi: [10.1186/s12888-016-0952-0](https://doi.org/10.1186/s12888-016-0952-0)] [Medline: [27400874](https://pubmed.ncbi.nlm.nih.gov/27400874/)]
36. De Choudhury M, Morris MR, White RW. *Seeking and sharing health information online: comparing search engines and social media*. New York, NY, USA: Association for Computing Machinery; 2014 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 26, 2014 - May 1, 2014; Toronto, Ontario, Canada p. 1365-1376. [doi: [10.1145/2556288.2557214](https://doi.org/10.1145/2556288.2557214)]
37. Lal S, Nguyen V, Theriault J. Seeking mental health information and support online: experiences and perspectives of young people receiving treatment for first-episode psychosis. *Early Interv Psychiatry* 2018 Jun;12(3):324-330. [doi: [10.1111/eip.12317](https://doi.org/10.1111/eip.12317)] [Medline: [26810026](https://pubmed.ncbi.nlm.nih.gov/26810026/)]
38. Yan Z, Wang T, Chen Y, Zhang H. Knowledge sharing in online health communities: A social exchange theory perspective. *Information and Management* 2016 Jul;53(5):643-653. [doi: [10.1016/j.im.2016.02.001](https://doi.org/10.1016/j.im.2016.02.001)]

Edited by G Eysenbach; submitted 09.07.21; peer-reviewed by R Kessler, C Kaylor-Hughes; comments to author 08.08.21; revised version received 03.10.21; accepted 03.10.21; published 11.11.21.

Please cite as:

Choi B, Kim H, Huh-Yoo J

Seeking Mental Health Support Among College Students in Video-Based Social Media: Content and Statistical Analysis of YouTube Videos

JMIR Form Res 2021;5(11):e31944

URL: <https://formative.jmir.org/2021/11/e31944>

doi: [10.2196/31944](https://doi.org/10.2196/31944)

PMID: [34762060](https://pubmed.ncbi.nlm.nih.gov/34762060/)

©Bogeum Choi, Heejun Kim, Jina Huh-Yoo. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.11.2021. 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

Formative Study of Mobile Phone Use for Family Planning Among Young People in Sierra Leone: Global Systematic Survey

Emeka Chukwu^{1,2}, BEng, MSc, CGEIT; Sonia Gilroy², BA, MA; Kojo Addaquay², BA; Nki Nafisa Jones², BA; Victor Gbadia Karimu³, BA, MA; Lalit Garg¹, BE, PGDIT, PhD; Kim Eva Dickson², BSc, MBChB, DFFP

¹Department of Computer Information System, Faculty of Information and Communications Technology (ICT), Msida, Malta

²United Nations Population Fund Country Office, Freetown, Sierra Leone

³National Secretariat for the Reduction of Teenage Pregnancy, New England, Freetown, Sierra Leone

Corresponding Author:

Emeka Chukwu, BEng, MSc, CGEIT

Department of Computer Information System

Faculty of Information and Communications Technology (ICT)

University of Malta

Msida, MSD 2080

Malta

Phone: 356 99330888

Email: nnaemeka_ec@hotmail.com

Abstract

Background: Teenage pregnancy remains high with low contraceptive prevalence among adolescents (aged 15-19 years) in Sierra Leone. Stakeholders leverage multiple strategies to address the challenge. Mobile technology is pervasive and presents an opportunity to reach young people with critical sexual reproductive health and family planning messages.

Objective: The objectives of this research study are to understand how mobile health (mHealth) is used for family planning, understand phone use habits among young people in Sierra Leone, and recommend strategies for mobile-enabled dissemination of family planning information at scale.

Methods: This formative research study was conducted using a systematic literature review and focus group discussions (FGDs). The literature survey assessed similar but existing interventions through a systematic search of 6 scholarly databases. Cross-sections of young people of both sexes and their support groups were engaged in 9 FGDs in an urban and a rural district in Sierra Leone. The FGD data were qualitatively analyzed using MAXQDA software (VERBI Software GmbH) to determine appropriate technology channels, content, and format for different user segments.

Results: Our systematic search results were categorized using Grading of Recommended Assessment and Evaluation (GRADE) into communication channels, audiovisual messaging format, purpose of the intervention, and message direction. The majority of reviewed articles report on SMS-based interventions. At the same time, most intervention purposes are for awareness and as helpful resources. Our survey did not find documented use of custom mHealth apps for family planning information dissemination. From the FGDs, more young people in Sierra Leone own basic mobile phones than those that have feature capabilities or are smartphone. Young people with smartphones use them mostly for WhatsApp and Facebook. Young people widely subscribe to the social media-only internet bundle, with the cost ranging from 1000 leones (US \$0.11) to 1500 leones (US \$0.16) daily. Pupils in both districts top-up their voice call and SMS credit every day between 1000 leones (US \$0.11) and 5000 leones (US \$0.52).

Conclusions: mHealth has facilitated family planning information dissemination for demand creation around the world. Despite the widespread use of social and new media, SMS is the scalable channel to reach literate and semiliterate young people. We have cataloged mHealth for contraceptive research to show SMS followed by call center as widely used channels. Jingles are popular for audiovisual message formats, mostly delivered as either push or pull only message directions (not both). Interactive voice response and automated calls are best suited to reach nonliterate young people at scale.

(*JMIR Form Res* 2021;5(11):e23874) doi:[10.2196/23874](https://doi.org/10.2196/23874)

KEYWORDS

young people; short message service; SMS; chatbot; text message; interactive voice response; IVR; WhatsApp; Facebook; family planning; contraceptives; Sierra Leone

Introduction

Sierra Leonean Context

Sierra Leone has a young population with an estimated 39.4% aged 15 to 35 years [1]. Pregnancy rates are high among adolescents. According to the 2013 Sierra Leone Demographic and Health Survey, overall 28% of adolescents aged 15 to 19 years have begun childbearing [2]. Among adolescents aged 15 to 19 years with no education, 46% have already started childbearing, which is more than double compared with 22% of those with secondary and higher education [2]. In Sierra Leone, cultural norms, low literacy, and limited access to information make young people prone to misinformation and inadequate knowledge of their sexual reproductive health and rights. Adolescents and young people hold many common myths and misconceptions about sexual reproductive health and family planning. A total of 36% of adolescents in the lowest wealth quintile have started childbearing compared to 14% in the highest wealth quintile [2].

Sierra Leone has low contraceptive use. Among the currently married, contraceptive use in young women aged 15 to 19 years is 8% [3]. Overall, urban dwellers have a high contraceptive prevalence of 26.6% compared with rural dwellers, with a prevalence rate of 13%. Myths, misinformation, and long-held traditional beliefs about contraceptives and long-term side effects on health, including infertility, are pervasive. The government is using different approaches to address these barriers, including awareness creation. Traditional contraceptive information dissemination methods use multichannel and engagement schemes as detailed in the national strategy to reduce adolescent pregnancy [4].

Funding from the United Nations Population Fund with the United Kingdom Department for International Development supports Sierra Leone's Government in developing and deploying a mobile-enabled family planning information awareness program. Researchers are increasingly using formative research to understand the best entry strategy, aid planning, and design interventions. Similarly, qualitative research is accepted in health programs and research [5] and is mainly used in formative research. Digital health, "the systematic application of information and communications technologies, computer science, and data to support informed decision-making by individuals, the health workforce, and health systems to strengthen resilience to disease and improve health and wellness" [6], often benefits from formative research-based design.

Mobile Technology

Mobile technology (mHealth) can democratize access to critical sexual reproductive health and family planning information by

reducing barriers such as stigma and fear often experienced by young people. Digital-enabled awareness can help eliminate traditional economic, geographic, and literacy constraints. Evidence shows that mobile technology, when appropriately applied, can help improve knowledge and awareness of end users [7] and caregivers [8]. Mobile phone ownership is emerging as a measure of socioeconomic status for underserved regions of the world [9]. Evidence from Kenya, with similar socioeconomic demographics as Sierra Leone, shows that mobile phone ownership is directly proportional to educational status, wealth, and having fewer children [9]. In Sierra Leone, technology is playing a pervasive role in several sectors, including the health care sector. Sierra Leone has an estimated 79 connected mobile SIM per 100 population teledensity [10]. The global pandemic occasioned by COVID-19 has shown the impact of technology in general and mHealth, particularly for information and services [11]. Emerging technologies beyond mHealth are increasingly being adopted for health and social service delivery [11]. There is an increasing surge in mobile apps for service delivery, particularly maternal and child health [12]. Similarly, policy makers are prioritizing digital technologies for strategic health care improvements [13].

Study Objective

This study is formative research to aid the design of a technologically appropriate and culturally relevant mHealth intervention in Sierra Leone. The specific objectives of the formative research were as follows:

- Understand the global trends on the mobile phone for family planning information dissemination
- Explore how young people in Sierra Leone use mobile phones
- Understand the needs and barriers to family planning information and services
- Gain information to support the design of a technology appropriate and culturally relevant mHealth intervention

The sections that follow detail the approach, main findings, and study recommendations.

Methods

This formative study involved a systematic scholarly literature search and review and focus group discussions (FGD).

Literature Review

We systematically searched 6 databases considered representative of both family planning (contraceptives) and mobile technology. We conducted a systematic search including relevant published papers from 2000 to 2019. We used the search keywords in [Table 1](#) and their extensions.

Table 1. Search terms and returned results.

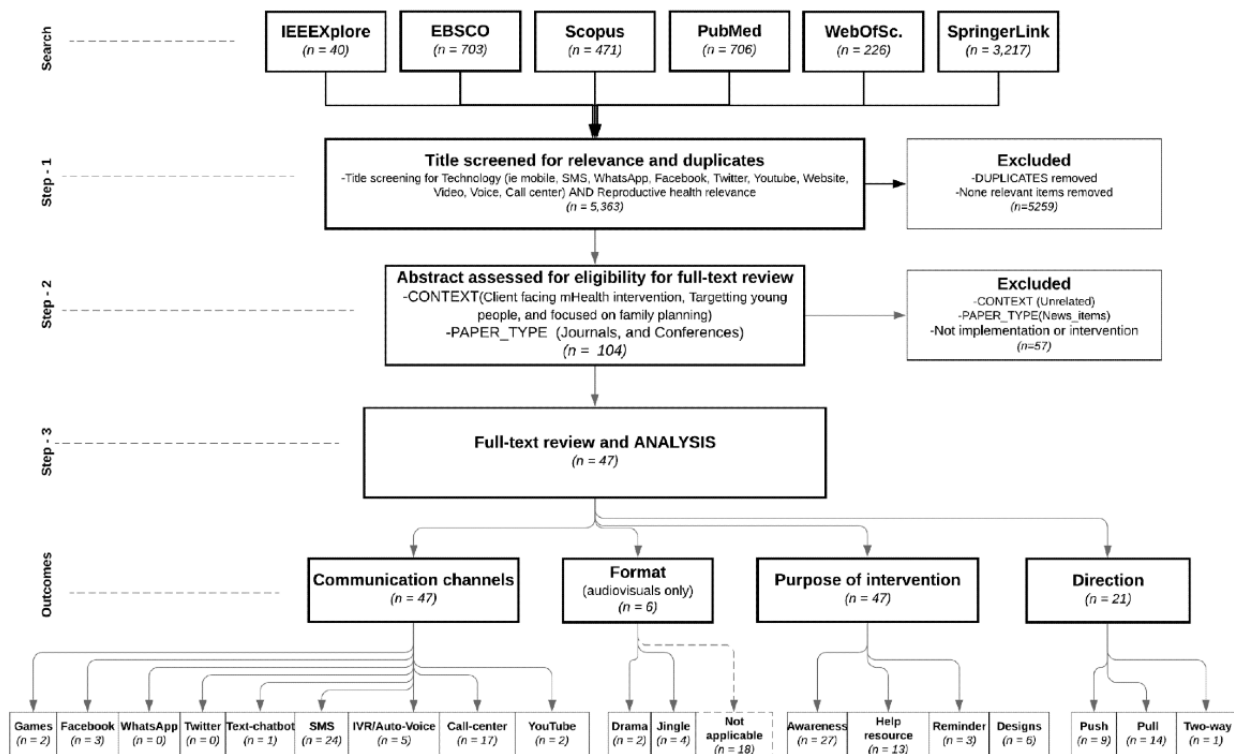
Search terms	Scopus	IEEE ^a	EBSCO	PubMed	Springer-Link	Web of Science	Total
SMS AND family planning	35	2	126	36	283	25	507
Chatbot AND family planning	2	0	1	0	6	0	9
Interactive voice response AND family planning	7	3	4	6	55	5	80
mHealth AND family planning	66	3	179	144	141	60	593
WhatsApp AND family planning	1	1	1	2	44	1	50
Facebook AND family planning	24	7	25	14	371	32	473
Twitter AND family planning	10	4	20	9	200	8	251
Hotline AND family planning	31	1	28	59	257	12	388
Call center AND family planning	124	15	9	114	106	1	369
YouTube AND family planning	7	1	1	6	172	0	187
SMS AND contraceptive	23	1	113	30	386	15	568
Chatbot AND contraceptive	1	0	1	0	3	0	5
Interactive voice response AND contraceptive	7	0	6	7	63	5	88
mHealth AND contraceptive	44	2	127	74	104	34	385
WhatsApp AND contraceptive	1	0	2	0	36	1	40
Facebook AND contraceptive	19	0	60	24	345	13	461
Twitter AND contraceptive	3	0	10	11	195	2	221
Hotline AND contraceptive	19	0	31	71	211	5	337
Call center AND contraceptive	45	0	11	94	78	2	230
YouTube AND contraceptive	2	0	10	5	161	5	183
Total	471	40	703	706	3217	226	5363

^aIEEE: Institute of Electrical and Electronics Engineers.

The search identified 5363 papers. After deduplication and removal of nonrelevant items through title screening, 104 papers remained. Titles were screened for inclusion if they reflected a combination of both technology and reproductive health. For technology, the keywords of interest were mobile, SMS, WhatsApp, Facebook, Twitter, YouTube, website, video, voice, call, call center, or their app synonyms. All returned results were included in an Excel (Microsoft Corp) research information template through this process. The 104 article abstracts were

then reviewed for context and paper type. Our context of interest was citizen-facing mHealth interventions for family planning. Paper types included were conference publications and journal articles. Only 47 papers were included after screening the 104 abstracts and excluding 57 publications for no implementation or not being related to our context. The number of items deemed eligible for full-text review came to 47 papers. See [Figure 1](#) for our PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) process to review the 47 papers.

Figure 1. PRISMA flowchart on our systematic review of global mobile health research focused on family planning.



We used the Grading of Recommended Assessment and Evaluation (GRADE) scheme to categorize the outcome of our full-text review and analysis into communication channels, audiovisual content format, intervention purpose, and message flow direction [14].

Focus Group Discussions

To better understand and engage stakeholders and end users (young people) who may be affected or impacted or who may influence mHealth intervention outcomes, we conducted a series of FGDs. The interactions helped gather feedback on how young people use mobile phones and their internet use habits, phone ownership, and expense patterns.

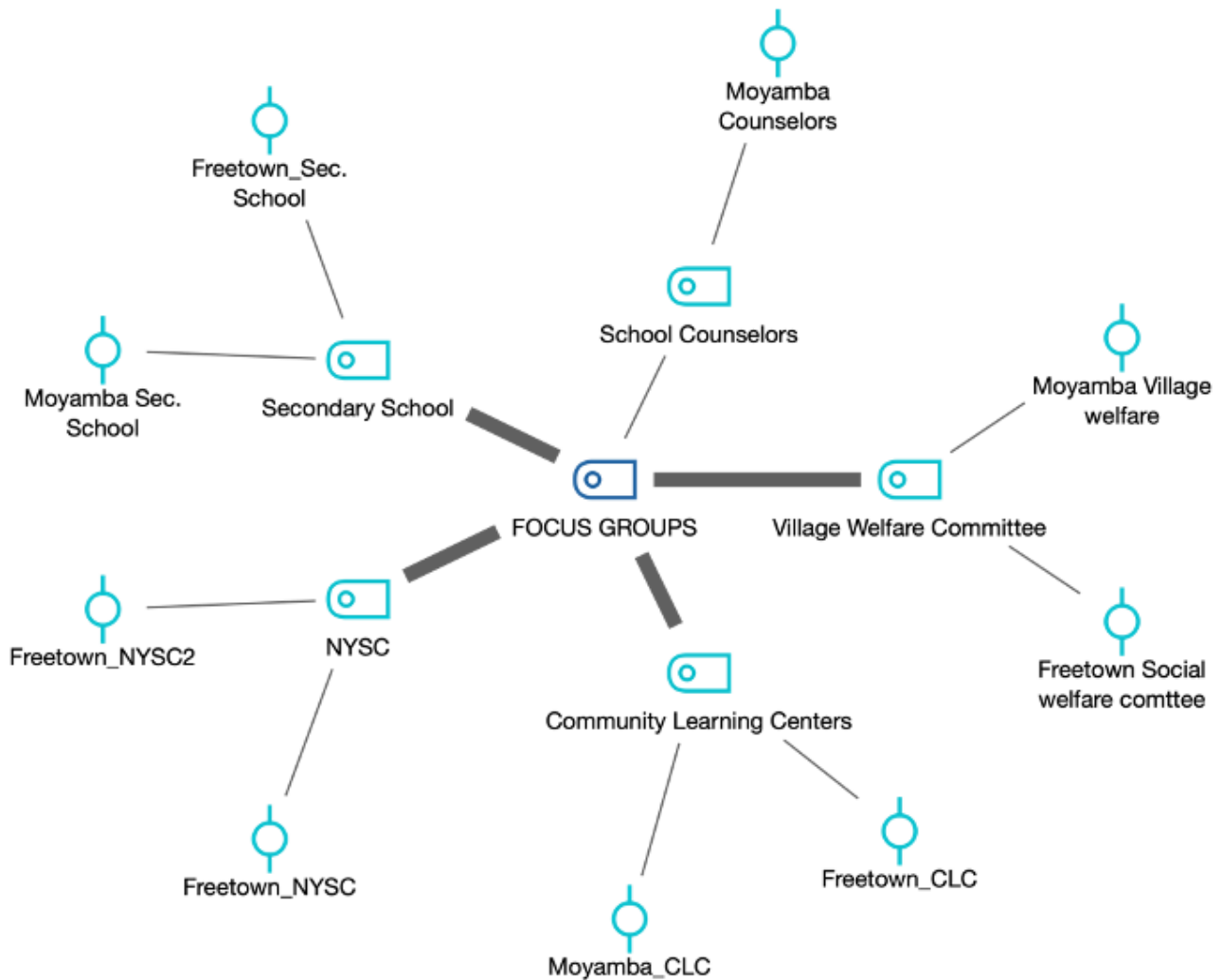
Participant Recruitment

We conducted the FGDs in 2 districts, one predominantly urban (ie, Freetown–Western Area Urban) and the other mostly rural (Moyamba). Moyamba district also has high early marriage with a median age of marriage being 17.5 years, below the national average, 18.2 years, and below the age of consent [15]. Also, 36.8% of girls in Moyamba have begun childbearing by age 18 years [16]. This is high compared to overall, with 28% of adolescents aged 15 to 19 years having begun childbearing. Target participants in both districts were purposely selected for each group targeting 5 to 10 participants in each group. Within each district, 4 groups were targeted as follows:

- Junior and senior secondary school students (male and female) to target end users with basic literacy, aged 14 to 18 years
- Youth corps members (male and female) to target end users with postsecondary education, aged 19 to 24 years
- Community learning center (CLC) participants to target females aged 14 to 24 years for nonliterate end users
- Village welfare committees, consisting of community members with authoritative knowledge of young people in the community

CLCs are centers where out-of-school girls learn and participate in vocational activities. Conversely, the village welfare committee is a variable member committee that oversees the community members’ welfare. Their work spans health, education, and environmental welfare issues, and they generally represent the parents and gatekeepers for young people in their communities. The National Youth Service Corps (NYSC) group are graduates undergoing a 1-year national service. The group of school counselors provides counseling services to secondary schools. In all, we conducted 9 FGDs with target participants from NYSC, CLC, secondary school, village welfare, and school counselors, as shown in Figure 2. Table 2 further illustrates the number, distribution, and characteristics of participants in the 9 FGDs in Freetown and Moyamba districts. The semistructured discussion guide was pretested with 5 participants: a nurse, 2 secondary school pupils, and 2 college students prior to the FDGs.

Figure 2. Focus group segments.



FGD Approach

Discussions were conducted in the Krio language for all 9 groups. Each session started with an introduction to the purpose of the visit and discussion using the research information template as in the guide introduction. Informed consent was sought from participants, each having a copy of the consent

form and the facilitator reading the text in the Krio language. Each consenting participant was required to sign a consent form without providing personal information. The transcripts were transcribed following key themes in the semistructured questionnaire immediately after each group session. The only quantitative data are the aggregate number of participants and their aggregate demographics.

Table 2. Characteristics of focus group discussion participants.

Days 2019	Discussion sites	Participants	Gender	Age (years)	District
Aug 19	Pretesting FGD ^a tool	5	Mixed	23-25	Western Urban
Aug 21	Secondary school (junior and senior)	7	Mixed	18-24	Western Urban
Aug 22	National Youth Service Corps	4	All female	>18	Western Urban
Aug 21	Community learning center	10	All female	14-19	Western Urban
Aug 22	Village welfare committee	7	Mixed	>23	Western Urban
Aug 29	Secondary school (junior and senior)	10	Mixed	14-19	Moyamba
Aug 28	Community learning center	8	All female	14-23	Moyamba
Aug 29	Village welfare committee and nurses	10	Mixed	>24	Moyamba
Aug 28	School counselors	11	Mixed	>24	Moyamba
Sep 4	National Youth Service Corps	10	Mixed	>18	Western Urban

^aFGD: focus group discussion.

Data Coding and Analysis

We identified and organized the FGD transcripts into key themes and coded using MAXQDA Analytics Pro (version 20.2, VERBI Software GmbH) [17]. The transcripts were coded by FGD, demography, phone ownership, phone type, phone use, phone top-up, phone expense, internet subscription, internet use, sexual and reproductive health (SRH) information access, language preference, current SRH medium, preferred SRH medium, comfort discussing SRH, and preferred SRH content. We then performed grouping using constant comparative analysis of these themes adapted from grounded theory proposed by Chun Tie et al [18].

Ethics Approval

Approval was sought and obtained to conduct the study from Sierra Leone's Ministry of Health and Sanitation's Ethics and Scientific Review Committee on August 13, 2019. We did not collect individually identifiable information during the FGDs.

Results

This section reveals the study findings and builds on existing global research frameworks and insights from Gonsalves et al [19] and McCarthy et al [20]. Findings from our systematic literature search are presented first followed by results from the FGDs.

Literature Review Findings: The Global State of Mobile for Family Planning

Based on our GRADE evaluation, reviewed papers' intervention channels include games, Facebook, WhatsApp, chatbots, SMS, interactive voice response (IVR), automated voice messages, call centers, and YouTube, as illustrated in Table 3. Call center and SMS are by far the most used channels from our findings. No study intervention used custom app, WhatsApp, or Twitter as an intervention channel. Similarly, the purpose of the majority of these interventions in Table 3 is for contraceptive awareness, followed by those providing resources for contraceptive information and service. A few were for reminders, design, or research purposes.

Table 3. Article distribution showing channels and purpose of intervention.

Channel type	Games	Facebook	WhatsApp	Twitter	Chatbots	SMS	IVR ^a /audiovoice	Call center	YouTube	Video
Purpose of intervention	— ^b	—	—	—	—	—	—	—	—	—
Awareness	[21,22]	[23,24]	—	—	—	[20,23,25-39]	[37,40-42]	[35,36,43-45]	[32,46]	[32]
Reminders	—	—	—	—	—	[47,48]	[40]	—	—	—
Help resource	[22]	—	—	—	[49]	[30,49,50]	[42]	[42,51-58]	—	—
Feasibility and design study	—	[59]	—	—	—	[60-62]	[58]	[63]	—	—

^aIVR: interactive voice response.

^bNot applicable.

IVR mobile content delivery channel has been used for maternal health [64], postabortion care [40], and family planning [58]. The earliest was conducted in Cambodia in 2013 [40]. Although most interventions did not indicate the language, those that did

were IVR-delivered contents, and the language of content delivery was the local language (not English). SMS interventions for family planning improved consumer knowledge by 14% (from 9.9% to 18.2%) compared to a control group with limited

knowledge [27]. No SMS intervention for family planning studies discussed the language of delivery. Social media has also been used in other regions for family planning-based demand generation like the peer-led safe sex Facebook group in China [59]. Other demand generation interventions include serious games to enhance sexuality education for young adolescents in Hong Kong [21]. Serious games use virtual reality-enabled games with engaging family planning information. A video-based mobile technology intervention has equally shown promise among adolescents in the United States [32]. In Kenya, the Shujaaz multimedia platform used various channels ranging from comic radio programs, a Facebook

campaign, and SMS [23]. Most of the papers reviewed did not provide details on the content structure.

Given the complexity in designing audiovisual interventions, we further analyzed and categorized the audiovisual intervention papers into message format and message direction. The SMS intervention similarly has a 160-character limit with directions push, pull, or 2-way interactive. Audiovisual messages were delivered in drama or jingle style as shown in Table 4. Call centers were mostly configured so that people call in. Some interventions used multiple channels [32,35-37], while some were designed for more than one purpose. The intervention by Smith et al [40], for instance, is for awareness and reminders, as shown in Table 3.

Table 4. Article distribution showing audiovisual formats and message direction.

	Message format		Message direction		
	Drama style	Jingle style	Pushed to recipient	Pulled by recipient	Two-way
IVR ^a /autovoice	[58]	[37,40-42]	[37,40-42]	— ^b	[58]
Call center	—	—	[35,42,48,51,53]	[36,40-45,52,54-57,63]	—
YouTube and videos	[32]	—	—	[32]	—

^aIVR: interactive voice response.

^bNot applicable.

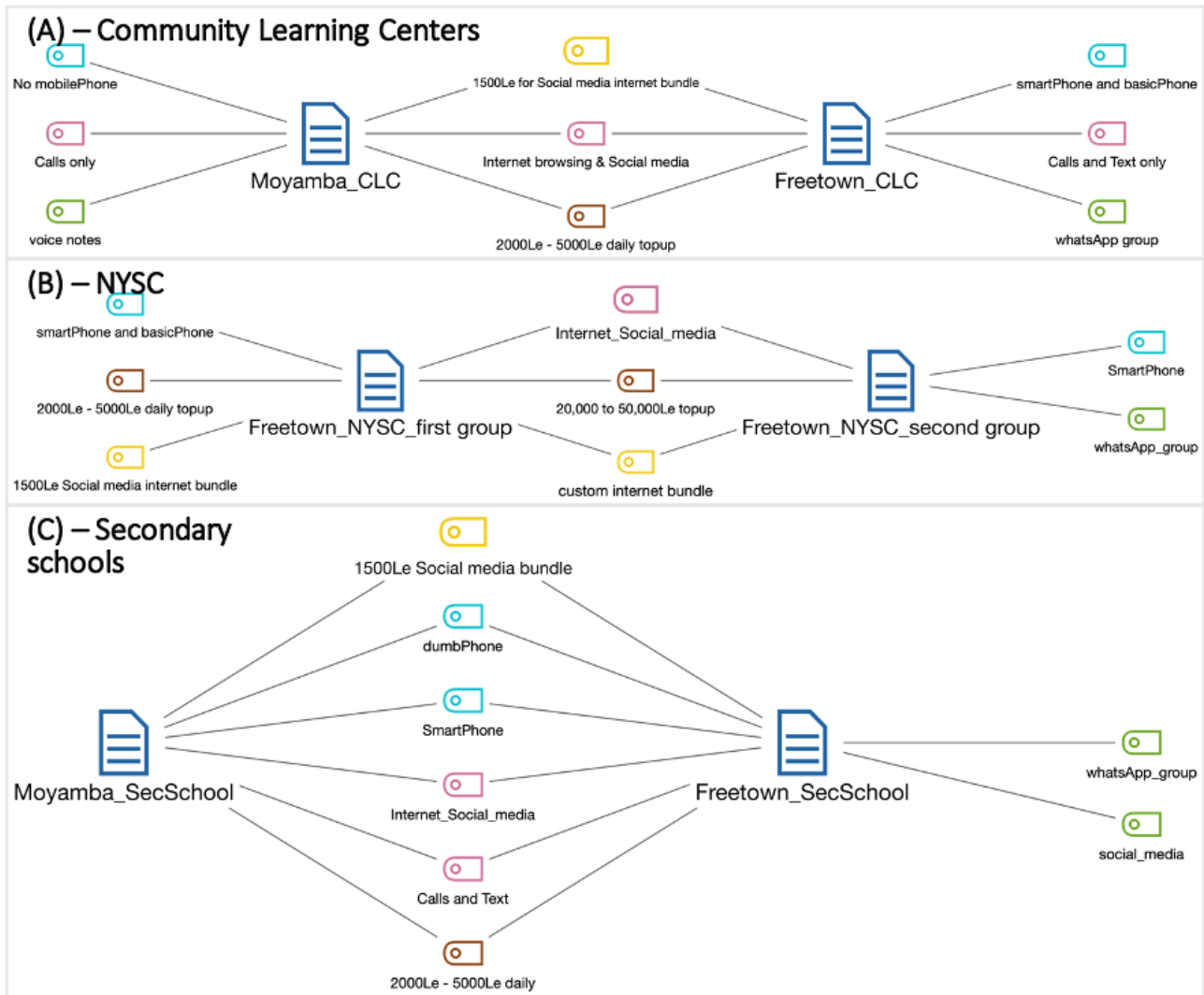
FGD Findings

Mobile Phone Ownership, Subscription, and Use Patterns

No participants in the CLC group in the Moyamba rural district had a mobile phone, although they indicated sharing with their family members (see blue in Figure 3A). The CLC group in the

Freetown urban district comprises participants who own smartphones and basic mobile phones (blue). Those who have used smartphones indicated subscribing exclusively to the social media bundle (yellow) among the CLC groups. Figure 3A shows these; both CLC groups credit or top-up their phones with 5000 leones (US \$0.50) or less to make calls (brown). The voice call credit (top-up) subscriptions were intermittent and far between for the CLC participants, especially those with basic phones.

Figure 3. Phone ownership, use, top-up, internet subscription, and internet use by focus group.



I sell top-up, and young people mostly buy 2000 leones.... Out of every 10 purchases, 8 of them will be 2000 leones. [Village welfare committee respondent]

Conversely, as shown in Figure 3B, all participants in the NYSC group had phones. In the second NYSC group, all participants had smartphones, most participants subscribed to a higher internet plan, and only one participant subscribed to the social media bundle. The majority of young people received their phones as gifts from parents, relatives, or friends.

Some girls can get their phone from boyfriend. [Moyamba district respondent]

I bought mine from a shop to pay back in installments. [Moyamba district respondent]

I can't use social media bundle because one can't do anything with it except Facebook and WhatsApp. [NYSC respondent]

You cannot even watch YouTube videos with the social media bundle. [NYSC respondent]

The NYSC participants' daily call credit use was less consistent but much higher than most other groups.

I will credit an average of 20,000 leones [US \$2.10] to 50,000 leones [US \$5.24] every now and then, and refill when exhausted. [NYSC respondent]

[I] refill with 4000 to 5000 leones [US \$0.42 to \$0.52] daily. [NYSC respondent]

I use unlimited internet plan that cost 500,000 leones [US \$55.56] because of my business. That is what works for me because other internet plans exhaust before the end of the month. [NYSC respondent]

Figure 3C illustrates the case for secondary school pupils.

Mobile Phone Use and Language

Participants in all groups were unanimous that WhatsApp group discussion and texts happen mostly in English, while video and audio messages or discussions are conducted mainly in Krio.

I only call with my phone and we speak Mende or Krio language... [Moyamba CLC respondent]

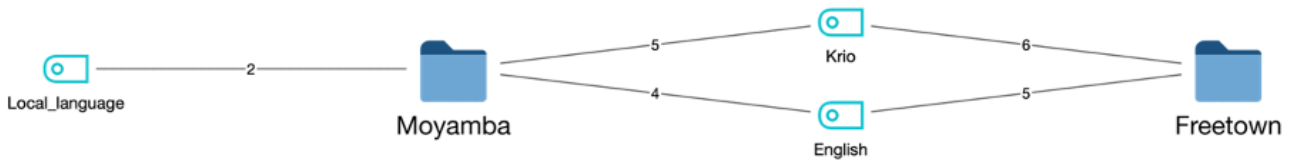
Calls are almost exclusively in Krio or local dialect. When asked if she sends SMS with her phone, she responded, "No."

Until you understand and write English, you cannot understand and write Krio. [School counselor respondent in the village welfare committee]

The qualitative analysis of responses from both districts, as shown in Figure 4, indicates that young people in Freetown preferred either English or Krio language, while those in Moyamba sometimes prefer any of their local dialects (Mende

or Temne) in addition to English and Krio. The FGDs revealed that the NYSC group uses WhatsApp the most, followed by Facebook and then banking or mobile money apps.

Figure 4. Language preference distribution.

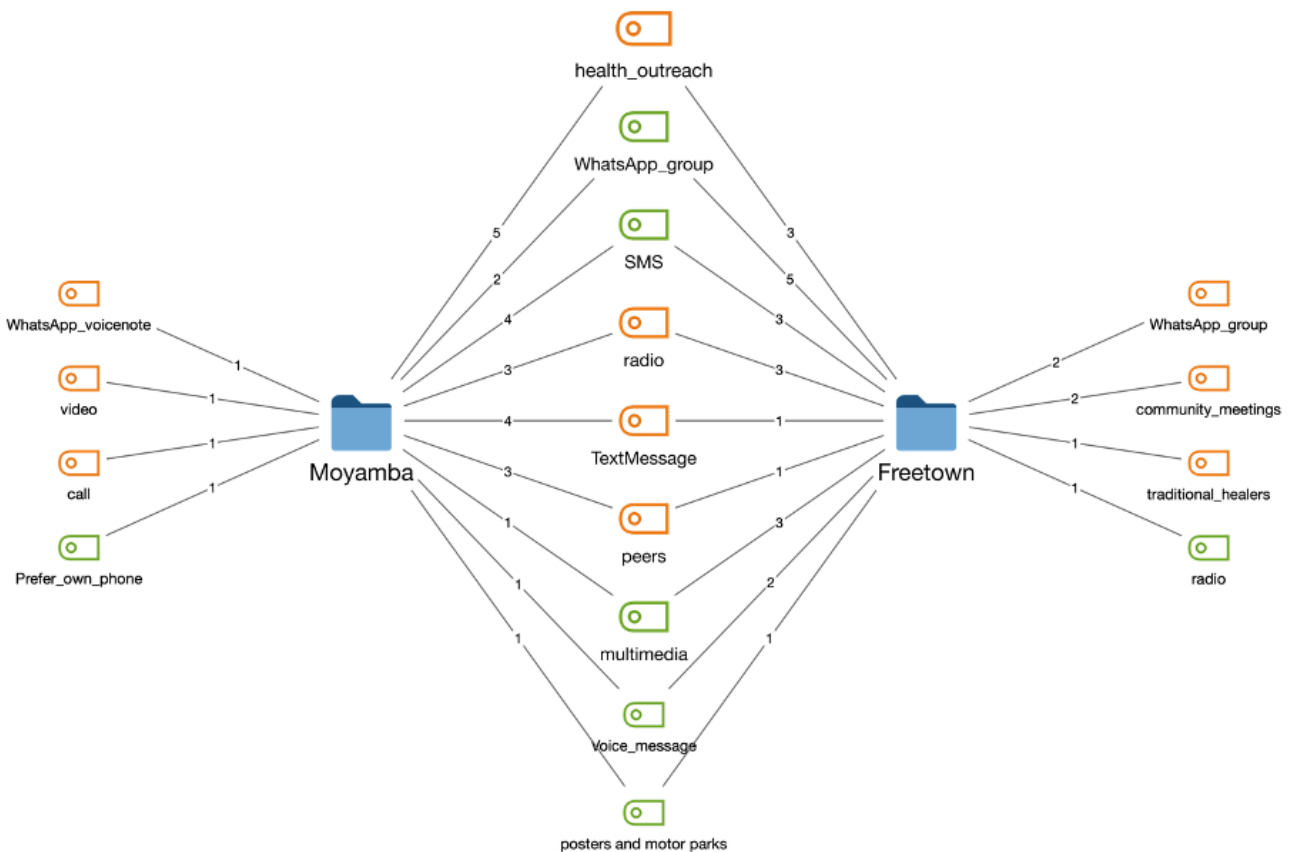


Young People’s Information Habits

In Figure 5, the color amber is for the current information medium, and the color green is for the preferred information medium based on focus group response analysis. The weighting

on the lines indicates how many groups indicated each option from the transcript. WhatsApp group, SMS, voice message, multimedia message, and posters at motor parks featured prominently.

Figure 5. Distribution of current and preferred sexual and reproductive health information medium by district.



People from Marie stopes always bring their buses with fitted speakers to parks and market places around to promote family planning. [Moyamba CLC respondent]

*I subscribed for daily health information for 2500 leones per week by dialing *931# on the Orange network. When there is no credit balance, the SMS messages are delayed and get delivered as soon as I top-up. [School counselor cohort respondent]*

Some of the health topics discussed include rape, early pregnancy, and early marriage. Peers are the medium

predominantly used for information clarification by secondary school pupils.

Awuko news sends regular health information from the Ministry of Health by SMS. [Moyamba CLC respondent]

A WhatsApp group is considered an effective avenue for message dissemination by this group. Secondary school pupils discuss their sexual health with their parents. Participants who own smartphones consistently showed a preference for WhatsApp as a message receipt channel compared to call or SMS channels.

I think SMS may be better for disseminating information compared to calls because you can refer to it again after reading the first time, and you can show others unlike calls. [NYSC respondent]

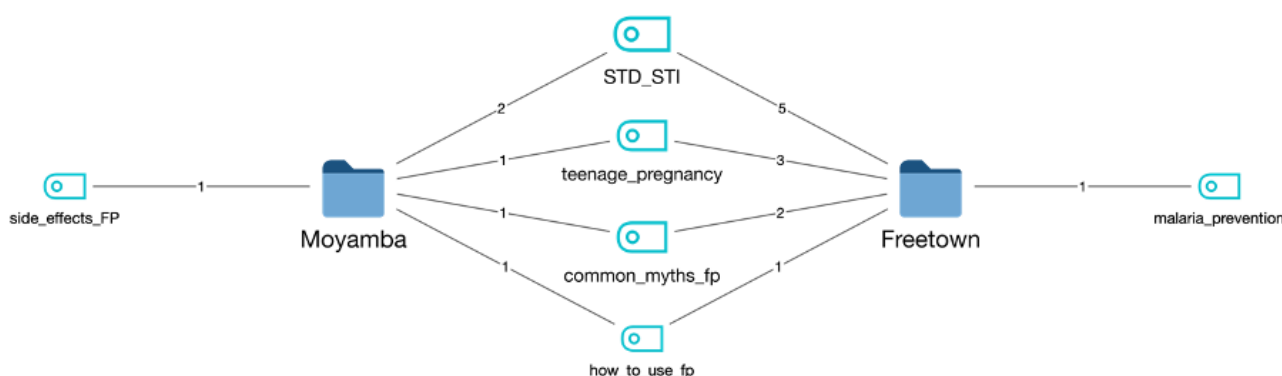
The NYSC group indicated they have more trust in messages received via SMS, although they could not explain why they think so.

SRH Information, Myths, Misconceptions, and Parents

Young people identified information on sexually transmitted diseases and sexually transmitted infections as what they want

to hear about the most. This is followed by a preference for information on how to prevent teenage pregnancy (see Figure 6). One issue identified was that parents and young people are worried about the long-term effect of contraceptives. A Moyamba CLC respondent told a story she heard from a friend about “a girl whose hand was cut off because the implant caused cancer.” When asked if she knew the girl in question, her response was “no.” Another issue that a nurse in the community welfare group tried to solve is one where “the parents force young girls to take out already inserted implants.”

Figure 6. Distribution of preferred sexual and reproductive health information medium by district.



The nurse explained that this is rampant and that health workers discourage this practice by increasing the cost of removal to 40,000 leones (US \$4.44). As a result of these reasons, some participants believe that some mothers fabricate myths to stop their girls from taking these contraceptives for fear of promiscuity.

Discussion

Our literature review, coupled with the FGDs, provided global and local insight and context to help us decide on an appropriate strategy for an mHealth for family planning intervention.

Understanding Global Trends in Mobile for Family Planning Information Dissemination

Our systematic review revealed health information had been delivered through mobile for maternal health [64], postabortion care [40], and family planning [27]. An SMS campaign in Kenya, for example, helped improve recipients' knowledge by 14 percentage points [27]. The 1-way pull or push message directions were predominant. Bidirectional intervention content was featured only once in the literature. Communication channels were Facebook, YouTube, SMS, voice calls, and call center. Most research study interventions were on SMS and call center, and none on WhatsApp or Twitter. We found from aggregated studies that the educational background, language, and phone ownership pattern directly correlate with the choice of a delivery channel. Only literate users, for instance, can access text-based content. Low-literate young people will be better reached through voice-based delivery like IVR and person-manned call centers. Low-literate users may not be reachable with multimedia like videos due to their often limited access to smart devices. Evidence of implementation or impact

of mobile demand generation initiatives for family planning is still lacking in Sierra Leone.

Related to our systematic survey, a study surveyed the 2 popular app stores for Android and iOS and found 5276 and 877 custom apps, respectively, designed for maternal and child health [12]. These apps did not appear in our systematic literature search of scholarly databases. There are 2 main reasons why this may be so: (1) they may not have met the study inclusion criteria or (2) they may not have had scholarly contributions at the time of survey and writing.

Exploring How Young People in Sierra Leone Use Mobile Phones

Mobile phone ownership appeared to increase with educational attainment, as a pattern seems to have emerged when we compared out-of-school young people at CLC or secondary school pupils or NYSC members. Ownership of a mobile phone is now considered an item of immense importance, and those who do not have one indicated sharing with their family and friends. To buttress this point, a participant told the following story.

My friends will always make fun of me for not having a phone. When I was opportune to volunteer for an NGO and was paid a stipend, I did not save any of the money. I went straight to the shop where I bought the android phone...to avoid the shame. [Freetown CLC respondent]

Despite this high interest, many young people still do not own their mobile phones, possibly due to funding constraints. Sharing mobile phones will often reduce the privacy of the content on that mobile. Participants who use or own smartphones prefer

WhatsApp and Facebook as the primary communication channels.

Conversely, basic mobile phone owners are constrained to performing essential phone functions: voice call and text messages. The majority of smartphone users outside the NYSC cohorts used the social media bundle, allowing Facebook and WhatsApp access only. This may limit how much rich multimedia content can be delivered to these young people. Participants with a smartphone use a social media bundle subscribed to daily. The expense pattern described indicates that even participants who subscribe to a social media bundle can miss out on any multimedia rich intervention even if they may be interested as a result of funding constraints. Options available will be to provide data to them for this or deliver content that is not dependent on internet data but available through the telecommunication network infrastructure.

Understanding the Need and Barriers to Family Planning Information and Services

The users at the bottom of the phone ownership pyramid understand Krio and their other local languages. These users do not generally write for communication. This means that those who cannot read in English cannot read in Krio, which corroborates the earlier finding. In deciding an intervention strategy, the user's phone type is only one of many factors to consider. There are other trade-offs to the final decision, including available resources, intended time to market, and technical implications. When users are asked without regard to their affordability and current spending habits, massive multimedia channels such as videos, social media, Mobi-sites, and mobile apps then become viable channels. Among the low-literate young people, there is a large proportion who still do not have a mobile, and they can be reached via radio and television.

Gaining Insights to Support the Design of Technologically Appropriate and Culturally Relevant mHealth Interventions

From the literature review and FGDs, it was determined that to reach young people who have basic phones and are less literate,

an audio-based system that would need to be delivered to a basic phone was necessary. The 3 options were either a call center, prerecorded automated calls, or an IVR. On the other hand, literate users with basic phones can be reached with text messages in the form of SMS. However, this would only reach a limited group, leaving out the most vulnerable. This has been proven to be effective in other countries [27]. Moreover, it is relatively less expensive and less technically involved once the message content has been developed.

Conclusion

The high incidence of teenage pregnancy and low contraceptive prevalence rate among young people necessitated the formative research. The study would help inform the design of an mHealth intervention as one component of broader efforts to delay sexual intercourse among young people and increase demand for family planning among those who are sexually active. The formative research helped show that globally, SMS followed by call center are the 2 widely used messaging channels. Audiovisual message formats were either drama or jingles delivered as push or pull only message directions (not both). For interventions that indicated message direction, only 1 of 21 implemented 2-way messaging; others were either push or pull 1-way messaging.

Solutions like automated SMS, call centers, IVR, chatbots, Facebook, and YouTube have been used worldwide. For scalability, the priority channels are SMS, call center, or IVR because they can reach users with basic mobile phones who make up most phone users among young people [10]. However, the volume of content delivered per time is limited for the same cost compared to other media-rich channels like videos, animations, and audio. Although not widely researched, multimedia channels like WhatsApp, Facebook, and video-based mobile apps are an emerging area appropriate and engaging for literate and affluent users. Young people's education level and socioeconomic disposition affect their mobile phone ownership, phone use patterns, and access to health information. These characteristics are relevant for designing appropriate mHealth interventions in Sierra Leone and elsewhere.

Acknowledgments

This formative study was funded through the UK Department for International Development's supported Saving Lives in Sierra Leone program.

Authors' Contributions

KED is the principal investigator for the study and reviewed and approved the different drafts of this work. EC is the coprincipal investigator and prepared the initial draft of this manuscript. SG provided technical input and reviewed the different drafts of the manuscript. KA supported the FGDs and reviewed the early drafts of the manuscript. NNJ facilitated the FGDs and reviewed the first draft of the manuscript. VGK cofacilitated the FGDs and reviewed the draft manuscript. LG also reviewed the draft manuscript. The authors are grateful to Dr Abiodun Oyeyipo for providing technical input and reviews to the manuscript drafts.

Conflicts of Interest

None declared.

References

1. Population by age and sex: Sierra Leone Demographic and Health Survey 2013. Statistics Sierra Leone. 2014. URL: <https://www.afro.who.int/sites/default/files/2017-06/SR215.pdf> [accessed 2021-01-09]
2. Teenage pregnancy and motherhood: Sierra Leone Demographic and Health Survey 2013. Statistics Sierra Leone. 2014. URL: <https://dhsprogram.com/pubs/pdf/SR215/SR215.pdf> [accessed 2021-01-12]
3. Current use of contraceptive methods: Sierra Leone Demographic and Health Survey 2013. Statistics Sierra Leone. 2014. URL: <https://www.afro.who.int/sites/default/files/2017-06/SR215.pdf> [accessed 2021-01-09]
4. National strategy for the reduction of adolescent pregnancy and child marriage 2018-2022. National Secretariat for the Reduction of Teenage Pregnancies. URL: https://nsrtp.sl/wp-content/uploads/2019/11/Sierra_Leone_National_Strategy_for_the_Reduction_of_Teenage_Pregnancy-2018.pdf [accessed 2020-12-18]
5. Eborall H, Morton K. Use of focus groups in developing behavioural mhealth interventions: a critical review. In: Barbour RS, Morgan DL, editors. *A New Era in Focus Group Research: Challenges, Innovation and Practice*. London: Palgrave MacMillan; 2017.
6. Digital implementation investment guide: integrating digital interventions into health programmes. Geneva: World Health Organization URL: <https://www.who.int/publications/i/item/9789240010567> [accessed 2021-01-09]
7. Green EP, Whitcomb A, Kahumbura C, Rosen JG, Goyal S, Achieng D, et al. "What is the best method of family planning for me?": a text mining analysis of messages between users and agents of a digital health service in Kenya. *Gates Open Res* 2019;3:1475 [FREE Full text] [doi: [10.12688/gatesopenres.12999.1](https://doi.org/10.12688/gatesopenres.12999.1)] [Medline: [31410395](https://pubmed.ncbi.nlm.nih.gov/31410395/)]
8. McNabb M, Chukwu E, Ojo O, Shekhar N, Gill CJ, Salami H, et al. Assessment of the quality of antenatal care services provided by health workers using a mobile phone decision support application in northern Nigeria: a pre/post-intervention study. *PLoS One* 2015;10(5):e0123940 [FREE Full text] [doi: [10.1371/journal.pone.0123940](https://doi.org/10.1371/journal.pone.0123940)] [Medline: [25942018](https://pubmed.ncbi.nlm.nih.gov/25942018/)]
9. Lee S, Begley CE, Morgan R, Chan W, Kim S. Addition of mHealth (mobile health) for family planning support in Kenya: disparities in access to mobile phones and associations with contraceptive knowledge and use. *Int Health* 2019 Nov 13;11(6):463-471. [doi: [10.1093/inthealth/ihy092](https://doi.org/10.1093/inthealth/ihy092)] [Medline: [30576546](https://pubmed.ncbi.nlm.nih.gov/30576546/)]
10. GSMA Mobile Connectivity Index. URL: <http://www.mobileconnectivityindex.com> [accessed 2020-05-12]
11. 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/)]
12. Matsuo S, Penneys NS. Glandular extramammary Paget's disease. *Arch Dermatol* 1985 Jun;121(6):716-717. [Medline: [2988460](https://pubmed.ncbi.nlm.nih.gov/2988460/)]
13. Sierra Leone national digital health strategy 2018-2023. URL: <https://mohs2017.files.wordpress.com/2019/02/sl-national-digital-health-strategy-nov-2018.pdf> [accessed 2020-07-20]
14. Dijkers M. Introducing GRADE: a systematic approach to rating evidence in systematic reviews and to guideline development. URL: https://ktdrr.org/products/update/v1n5/dijkers_grade_ktupdatev1n5.pdf [accessed 2020-12-20]
15. Age at first sexual intercourse: Sierra Leone Demographic and Health Survey 2013. Statistics Sierra Leone. 2014. URL: <https://www.afro.who.int/sites/default/files/2017-06/SR215.pdf> [accessed 2021-01-03]
16. Teenage pregnancy and motherhood: Sierra Leone Demographic and Health Survey 2013. Statistics Sierra Leone. 2014. URL: <https://dhsprogram.com/pubs/pdf/fr297/fr297.pdf> [accessed 2020-07-10]
17. Kuckartz U, Rädiker S. Introduction: analyzing qualitative data with software. In: *Analyzing Qualitative Data with MAXQDA*. Cham: Springer International Publishing; 2019:1-11.
18. Chun Tie Y, Birks M, Francis K. Grounded theory research: a design framework for novice researchers. *SAGE Open Med* 2019;7:2050312118822927 [FREE Full text] [doi: [10.1177/2050312118822927](https://doi.org/10.1177/2050312118822927)] [Medline: [30637106](https://pubmed.ncbi.nlm.nih.gov/30637106/)]
19. Gonsalves L, L'Engle KL, Tamrat T, Plourde KF, Mangone ER, Agarwal S, et al. Adolescent/Youth Reproductive Mobile Access and Delivery Initiative for Love and Life Outcomes (ARMADILLO) Study: formative protocol for mHealth platform development and piloting. *Reprod Health* 2015 Aug 07;12:67 [FREE Full text] [doi: [10.1186/s12978-015-0059-y](https://doi.org/10.1186/s12978-015-0059-y)] [Medline: [26248769](https://pubmed.ncbi.nlm.nih.gov/26248769/)]
20. McCarthy OL, Wazwaz O, Osorio Calderon V, Jado I, Saibov S, Stavridis A, et al. Development of an intervention delivered by mobile phone aimed at decreasing unintended pregnancy among young people in three lower middle income countries. *BMC Public Health* 2018 May 02;18(1):576 [FREE Full text] [doi: [10.1186/s12889-018-5477-7](https://doi.org/10.1186/s12889-018-5477-7)] [Medline: [29716571](https://pubmed.ncbi.nlm.nih.gov/29716571/)]
21. Enhancing sexuality education for young adolescents through serious gaming. 2014. URL: <http://hub.hku.hk/handle/10722/204569> [accessed 2021-01-12]
22. Chu SKW, Kwan ACM, Reynolds R, Mellecker RR, Tam F, Lee G, et al. Promoting sex education among teenagers through an interactive game: reasons for success and implications. *Games Health J* 2015 Jun;4(3):168-174. [doi: [10.1089/g4h.2014.0059](https://doi.org/10.1089/g4h.2014.0059)] [Medline: [26182060](https://pubmed.ncbi.nlm.nih.gov/26182060/)]
23. Hutchinson P, Mirzoyants A, Leyton A. Empowering youth for social change through the Shujaaz multimedia platform in Kenya. *Int J Adolesc Youth* 2018 Sep 15;24(1):102-116. [doi: [10.1080/02673843.2018.1475287](https://doi.org/10.1080/02673843.2018.1475287)]
24. Purdy CH. Using the Internet and social media to promote condom use in Turkey. *Reprod Health Matters* 2011 May;19(37):157-165. [doi: [10.1016/S0968-8080\(11\)37549-0](https://doi.org/10.1016/S0968-8080(11)37549-0)] [Medline: [21555096](https://pubmed.ncbi.nlm.nih.gov/21555096/)]

25. Green EP, Whitcomb A, Kahumbura C, Rosen JG, Goyal S, Achieng D, et al. "What is the best method of family planning for me?": a text mining analysis of messages between users and agents of a digital health service in Kenya. *Gates Open Res* 2019;3:1475 [FREE Full text] [doi: [10.12688/gatesopenres.12999.1](https://doi.org/10.12688/gatesopenres.12999.1)] [Medline: [31410395](https://pubmed.ncbi.nlm.nih.gov/31410395/)]
26. McCarthy OL, Zghayyer H, Stavridis A, Adada S, Ahamed I, Leurent B, et al. A randomized controlled trial of an intervention delivered by mobile phone text message to increase the acceptability of effective contraception among young women in Palestine. *Trials* 2019 Apr 23;20(1):228 [FREE Full text] [doi: [10.1186/s13063-019-3297-4](https://doi.org/10.1186/s13063-019-3297-4)] [Medline: [31014358](https://pubmed.ncbi.nlm.nih.gov/31014358/)]
27. Johnson D, Juras R, Riley P, Chatterji M, Sloane P, Choi SK, et al. A randomized controlled trial of the impact of a family planning mHealth service on knowledge and use of contraception. *Contraception* 2017 Jan;95(1):90-97. [doi: [10.1016/j.contraception.2016.07.009](https://doi.org/10.1016/j.contraception.2016.07.009)] [Medline: [27421767](https://pubmed.ncbi.nlm.nih.gov/27421767/)]
28. Harrington EK, McCoy EE, Drake AL, Matemo D, John-Stewart G, Kinuthia J, et al. Engaging men in an mHealth approach to support postpartum family planning among couples in Kenya: a qualitative study. *Reprod Health* 2019 Feb 11;16(1):17 [FREE Full text] [doi: [10.1186/s12978-019-0669-x](https://doi.org/10.1186/s12978-019-0669-x)] [Medline: [30744697](https://pubmed.ncbi.nlm.nih.gov/30744697/)]
29. Olsen PS, Plourde KF, Lasway C, van Praag E. Insights from a text messaging-based sexual and reproductive health information program in Tanzania (m4RH): retrospective analysis. *JMIR Mhealth Uhealth* 2018 Nov 01;6(11):e10190 [FREE Full text] [doi: [10.2196/10190](https://doi.org/10.2196/10190)] [Medline: [30389651](https://pubmed.ncbi.nlm.nih.gov/30389651/)]
30. Perrier T. Male partner engagement in family planning SMS conversations at Kenyan health clinics. *Proc 1st ACM SIGCAS Conf Comput Sustainable Soc (COMPASS)* 2018:1-11. [doi: [10.1145/3209811.3209857](https://doi.org/10.1145/3209811.3209857)]
31. Blanc AK, Glazer K, Ofomata-Aderemi U, Akinfaderin-Agarau F. *Stud Fam Plann* 2016 Mar;47(1):39-53. [doi: [10.1111/j.1728-4465.2016.00046.x](https://doi.org/10.1111/j.1728-4465.2016.00046.x)] [Medline: [26952714](https://pubmed.ncbi.nlm.nih.gov/26952714/)]
32. Walsh-Buhi ER, Helmy H, Harsch K, Rella N, Godcharles C, Ogunrunde A, et al. Pregnancy prevention at her fingertips: a text- and mobile video-based pilot intervention to promote contraceptive methods among college women. *Health Education Journal* 2016 Jul 28;75(5):621-636. [doi: [10.1177/0017896915621939](https://doi.org/10.1177/0017896915621939)]
33. Gonsalves L, Hindin MJ, Bayer A, Carcamo CP, Gichangi P, Habib N, et al. Protocol of an open, three-arm, individually randomized trial assessing the effect of delivering sexual and reproductive health information to young people (aged 13-24) in Kenya and Peru via mobile phones: adolescent/youth reproductive mobile access and delivery initiative for love and life outcomes (ARMADILLO) study stage 2. *Reprod Health* 2018 Jul 11;15(1):126 [FREE Full text] [doi: [10.1186/s12978-018-0568-6](https://doi.org/10.1186/s12978-018-0568-6)] [Medline: [29996854](https://pubmed.ncbi.nlm.nih.gov/29996854/)]
34. Ampt FH, Mudogo C, Gichangi P, Lim MSC, Manguro G, Chersich M, et al. WHISPER or SHOUT study: protocol of a cluster-randomised controlled trial assessing mHealth sexual reproductive health and nutrition interventions among female sex workers in Mombasa, Kenya. *BMJ Open* 2017 Aug 18;7(8):e017388 [FREE Full text] [doi: [10.1136/bmjopen-2017-017388](https://doi.org/10.1136/bmjopen-2017-017388)] [Medline: [28821530](https://pubmed.ncbi.nlm.nih.gov/28821530/)]
35. Castaño PM, Bynum JY, Andrés R, Lara M, Westhoff C. Effect of daily text messages on oral contraceptive continuation: a randomized controlled trial. *Obstet Gynecol* 2012 Jan;119(1):14-20. [doi: [10.1097/AOG.0b013e31823d4167](https://doi.org/10.1097/AOG.0b013e31823d4167)] [Medline: [22143257](https://pubmed.ncbi.nlm.nih.gov/22143257/)]
36. Vu LTH, Nguyen NTK, Tran HTD, Muhajarine N. mHealth information for migrants: an e-health intervention for internal migrants in Vietnam. *Reprod Health* 2016 May 14;13(1):55 [FREE Full text] [doi: [10.1186/s12978-016-0172-6](https://doi.org/10.1186/s12978-016-0172-6)] [Medline: [27180147](https://pubmed.ncbi.nlm.nih.gov/27180147/)]
37. Brody C, Tuot S, Chhoun P, Swendenman D, Kaplan KC, Yi S. Mobile Link: a theory-based messaging intervention for improving sexual and reproductive health of female entertainment workers in Cambodia: study protocol of a randomized controlled trial. *Trials* 2018 Apr 19;19(1):235. [doi: [10.1186/s13063-018-2614-7](https://doi.org/10.1186/s13063-018-2614-7)]
38. Ashcroft N, Shelus V, Garg H, McLarnon-Silk C, Jennings VH. Implementation of CycleTel Family Advice: an SMS-based service to provide family planning and fertility awareness information in India. *Mhealth* 2017;3:20 [FREE Full text] [doi: [10.21037/mhealth.2017.05.01](https://doi.org/10.21037/mhealth.2017.05.01)] [Medline: [28607906](https://pubmed.ncbi.nlm.nih.gov/28607906/)]
39. Vahdat HL, L'Engle KL, Plourde KF, Magaria L, Olawo A. There are some questions you may not ask in a clinic: providing contraception information to young people in Kenya using SMS. *Int J Gynaecol Obstet* 2013 Nov;123 Suppl 1:e2-e6 [FREE Full text] [doi: [10.1016/j.ijgo.2013.07.009](https://doi.org/10.1016/j.ijgo.2013.07.009)] [Medline: [24012514](https://pubmed.ncbi.nlm.nih.gov/24012514/)]
40. Smith C, Vannak U, Sokhey L, Ngo TD, Gold J, Khut K, et al. MOBILE Technology for Improved Family Planning Services (MOTIF): study protocol for a randomised controlled trial. *Trials* 2013 Dec 12;14:427 [FREE Full text] [doi: [10.1186/1745-6215-14-427](https://doi.org/10.1186/1745-6215-14-427)] [Medline: [24330763](https://pubmed.ncbi.nlm.nih.gov/24330763/)]
41. Reiss K, Andersen K, Pearson E, Biswas K, Taleb F, Ngo TD, et al. Unintended consequences of mhealth interactive voice messages promoting contraceptive use after menstrual regulation in bangladesh: intimate partner violence results from a randomized controlled trial. *Glob Health Sci Pract* 2019 Sep;7(3):386-403 [FREE Full text] [doi: [10.9745/GHSP-D-19-00015](https://doi.org/10.9745/GHSP-D-19-00015)] [Medline: [31558596](https://pubmed.ncbi.nlm.nih.gov/31558596/)]
42. Reiss K, Andersen K, Barnard S, Ngo TD, Biswas K, Smith C, et al. Using automated voice messages linked to telephone counselling to increase post-menstrual regulation contraceptive uptake and continuation in Bangladesh: study protocol for a randomised controlled trial. *BMC Public Health* 2017 Oct 03;17(1):769 [FREE Full text] [doi: [10.1186/s12889-017-4703-z](https://doi.org/10.1186/s12889-017-4703-z)] [Medline: [28974209](https://pubmed.ncbi.nlm.nih.gov/28974209/)]
43. Trussell J, Koenig J, Vaughan B, Stewart F. Evaluation of a media campaign to increase knowledge about emergency contraception. *Contraception* 2001 Feb;63(2):81-87. [doi: [10.1016/s0010-7824\(01\)00174-3](https://doi.org/10.1016/s0010-7824(01)00174-3)] [Medline: [11292472](https://pubmed.ncbi.nlm.nih.gov/11292472/)]

44. LaRoche KJ, Foster AM. Toll free but not judgment free: evaluating postabortion support services in Ontario. *Contraception* 2015 Nov;92(5):469-474. [doi: [10.1016/j.contraception.2015.08.003](https://doi.org/10.1016/j.contraception.2015.08.003)] [Medline: [26260687](https://pubmed.ncbi.nlm.nih.gov/26260687/)]
45. Trussell J, Bull J, Koenig J, Bass M, Allina A, Gamble VN. Call 1-888-NOT-2-LATE: promoting emergency contraception in the United States. *J Am Med Womens Assoc* (1972) 1998;53(5 Suppl 2):247-250. [Medline: [9859633](https://pubmed.ncbi.nlm.nih.gov/9859633/)]
46. Hong J, Tsai C, Fan-Chiang C, Hwang M. Mindfulness in learning safe sex via social media: perspectives of personality and experiential value. *Comput Human Behav* 2016 Nov;64:337-346. [doi: [10.1016/j.chb.2016.06.033](https://doi.org/10.1016/j.chb.2016.06.033)]
47. Sari F. Early warning system for the continuous use of injectable contraception, using sms gateway in the success of family planning programs. *ICSCEE* 2018:1-6. [doi: [10.1109/icscee.2018.8538407](https://doi.org/10.1109/icscee.2018.8538407)]
48. Bracken H, Lohr PA, Taylor J, Morroni C, Winikoff B. RU OK? The acceptability and feasibility of remote technologies for follow-up after early medical abortion. *Contraception* 2014 Jul;90(1):29-35. [doi: [10.1016/j.contraception.2014.03.016](https://doi.org/10.1016/j.contraception.2014.03.016)] [Medline: [24815098](https://pubmed.ncbi.nlm.nih.gov/24815098/)]
49. Hussain S, Ogundimu F, Bhattarai S. Mobile phone-based chatbot for family planning and contraceptive information. In: Duffy V, editor. *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management*. Cham: Springer; 2019:342-352.
50. Unger JA, Ronen K, Perrier T, DeRenzi B, Slyker J, Drake AL, et al. Short message service communication improves exclusive breastfeeding and early postpartum contraception in a low- to middle-income country setting: a randomised trial. *BJOG* 2018 Nov;125(12):1620-1629. [doi: [10.1111/1471-0528.15337](https://doi.org/10.1111/1471-0528.15337)] [Medline: [29924912](https://pubmed.ncbi.nlm.nih.gov/29924912/)]
51. Kulathinal S, Joseph B, Säävälä M. Mobile helpline and reversible contraception: lessons from a controlled before-and-after study in rural India. *JMIR Mhealth Uhealth* 2019 Aug 09;7(8):e12672 [FREE Full text] [doi: [10.2196/12672](https://doi.org/10.2196/12672)] [Medline: [31400103](https://pubmed.ncbi.nlm.nih.gov/31400103/)]
52. Messinger CJ, Mahmud I, Kanan S, Jahangir YT, Sarker M, Rashid SF. Utilization of mobile phones for accessing menstrual regulation services among low-income women in Bangladesh: a qualitative analysis. *Reprod Health* 2017 Jan 14;14(1):7 [FREE Full text] [doi: [10.1186/s12978-016-0274-1](https://doi.org/10.1186/s12978-016-0274-1)] [Medline: [28088232](https://pubmed.ncbi.nlm.nih.gov/28088232/)]
53. Gupta N, Gupta A, Bhagia S, Singh S. Mobile technology for increasing postpartum family planning acceptability: the development of a mobile-based (mHealth) intervention through a dedicated counselor—a pilot innovative study conducted in a tertiary teaching hospital of Agra, Uttar Pradesh, India. *J South Asian Fed Obstet Gynaecol* 2018;10(1):74-80. [doi: [10.5005/jp-journals-10006-1564](https://doi.org/10.5005/jp-journals-10006-1564)]
54. Steyn P, Britz A, Rolink H. A 24-hour toll-free hotline for youth emergency contraceptive services in the Western Cape Province, South Africa. *Contraception* 2005 Sep;72(3):229. [doi: [10.1016/j.contraception.2005.06.002](https://doi.org/10.1016/j.contraception.2005.06.002)]
55. Yagnik P, Gold J, Stoope M, Reichwein B, van Gemert C, Corby N, et al. Development and pilot of a framework to evaluate reproductive health call centre services: experience of Marie Stopes international. *BMC Health Serv Res* 2015 Sep 21;15:398 [FREE Full text] [doi: [10.1186/s12913-015-1064-0](https://doi.org/10.1186/s12913-015-1064-0)] [Medline: [26392363](https://pubmed.ncbi.nlm.nih.gov/26392363/)]
56. Dabbs K. Family planning hotline. *Fam Plann Perspect* 1970 Jan;2(1):13-14. [Medline: [5520496](https://pubmed.ncbi.nlm.nih.gov/5520496/)]
57. Soriano MB. "Instant sagot": hotline for advice on family planning. *Initiatives Popul* 1975 Sep;1(1):32-42. [Medline: [12277224](https://pubmed.ncbi.nlm.nih.gov/12277224/)]
58. Babalola S, Loehr C, Oyenubi O, Akiode A, Mobley A. Efficacy of a digital health tool on contraceptive ideation and use in Nigeria: results of a cluster-randomized control trial. *Glob Health Sci Pract* 2019 Jun;7(2):273-288 [FREE Full text] [doi: [10.9745/GHSP-D-19-00066](https://doi.org/10.9745/GHSP-D-19-00066)] [Medline: [31249023](https://pubmed.ncbi.nlm.nih.gov/31249023/)]
59. Sun WH, Wong CKH, Wong WCW. A peer-led, social media-delivered, safer sex intervention for Chinese college students: randomized controlled trial. *J Med Internet Res* 2017 Dec 09;19(8):e284 [FREE Full text] [doi: [10.2196/jmir.7403](https://doi.org/10.2196/jmir.7403)] [Medline: [28793980](https://pubmed.ncbi.nlm.nih.gov/28793980/)]
60. Green EP, Augustine A, Naanyu V, Hess A, Kiwinda L. Developing a digital marketplace for family planning: pilot randomized encouragement trial. *J Med Internet Res* 2018 Jul 31;20(7):e10756 [FREE Full text] [doi: [10.2196/10756](https://doi.org/10.2196/10756)] [Medline: [30064968](https://pubmed.ncbi.nlm.nih.gov/30064968/)]
61. L'Engle KL, Vahdat HL, Ndakidemi E, Lasway C, Zan T. Evaluating feasibility, reach and potential impact of a text message family planning information service in Tanzania. *Contraception* 2013 Feb;87(2):251-256. [doi: [10.1016/j.contraception.2012.07.009](https://doi.org/10.1016/j.contraception.2012.07.009)] [Medline: [22935322](https://pubmed.ncbi.nlm.nih.gov/22935322/)]
62. Lim MSC, Hocking JS, Aitken CK, Fairley CK, Jordan L, Lewis JA, et al. Impact of text and email messaging on the sexual health of young people: a randomised controlled trial. *J Epidemiol Community Health* 2012 Jan;66(1):69-74. [doi: [10.1136/jech.2009.100396](https://doi.org/10.1136/jech.2009.100396)] [Medline: [21415232](https://pubmed.ncbi.nlm.nih.gov/21415232/)]
63. Baxter D, Williams JI. An investigation into the quality of service provided by telephone hotlines for family planning services. *Can J Public Health* 1982;73(3):194-199. [Medline: [7127242](https://pubmed.ncbi.nlm.nih.gov/7127242/)]
64. Murthy N, Chandrasekharan S, Prakash MP, Kaonga NN, Peter J, Ganju A, et al. The impact of an mhealth voice message service (mMitra) on infant care knowledge, and practices among low-income women in India: findings from a pseudo-randomized controlled trial. *Matern Child Health J* 2019 Dec;23(12):1658-1669 [FREE Full text] [doi: [10.1007/s10995-019-02805-5](https://doi.org/10.1007/s10995-019-02805-5)] [Medline: [31584144](https://pubmed.ncbi.nlm.nih.gov/31584144/)]

Abbreviations

CLC: community learning center

FGD: focus group discussion

GRADE: Grading of Recommended Assessment and Evaluation

IVR: interactive voice response

mHealth: mobile health

NYSC: National Youth Service Corps

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analysis

SRH: sexual and reproductive health

Edited by G Eysenbach; submitted 26.08.20; peer-reviewed by P Mechael, J Ye, N Mohammad Gholi Mezerji, T Scherr; comments to author 06.10.20; revised version received 26.10.20; accepted 13.04.21; published 12.11.21.

Please cite as:

Chukwu E, Gilroy S, Addaquay K, Jones NN, Karimu VG, Garg L, Dickson KE

Formative Study of Mobile Phone Use for Family Planning Among Young People in Sierra Leone: Global Systematic Survey

JMIR Form Res 2021;5(11):e23874

URL: <https://formative.jmir.org/2021/11/e23874>

doi: [10.2196/23874](https://doi.org/10.2196/23874)

PMID: [34766908](https://pubmed.ncbi.nlm.nih.gov/34766908/)

©Emeka Chukwu, Sonia Gilroy, Kojo Addaquay, Nki Nafisa Jones, Victor Gbadia Karimu, Lalit Garg, Kim Eva Dickson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.11.2021. 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

Dropout From an Internet-Delivered Cognitive Behavioral Therapy Intervention for Adults With Depression and Anxiety: Qualitative Study

Kate Lawler^{1,2}, BA, MSc; Caroline Earley^{1,2}, MSc; Ladislav Timulak¹, PhD; Angel Enrique^{1,2}, PhD; Derek Richards^{1,2}, PhD

¹E-Mental Health Research Group, School of Psychology, Trinity College Dublin, Dublin, Ireland

²Clinical Research and Innovation, SilverCloud Health, Dublin, Ireland

Corresponding Author:

Kate Lawler, BA, MSc
E-Mental Health Research Group
School of Psychology
Trinity College Dublin
College Green
Dublin
Ireland
Phone: 353 851510008
Email: lawlerka@tcd.ie

Abstract

Background: Treatment dropout continues to be reported from internet-delivered cognitive behavioral therapy (iCBT) interventions, and lower completion rates are generally associated with lower treatment effect sizes. However, evidence is emerging to suggest that completion of a predefined number of modules is not always necessary for clinical benefit or consideration of the needs of each individual patient.

Objective: The aim of this study is to perform a qualitative analysis of patients' experiences with an iCBT intervention in a routine care setting to achieve a deeper insight into the phenomenon of dropout.

Methods: A total of 15 purposively sampled participants (female: 8/15, 53%) from a larger parent randomized controlled trial were interviewed via telephone using a semistructured interview schedule that was developed based on the existing literature and research on dropout in iCBT. Data were analyzed using a descriptive-interpretive approach.

Results: The experience of treatment leading to dropout can be understood in terms of 10 domains: relationship to technology, motivation to start, background knowledge and attitudes toward iCBT, perceived change in motivation, usage of the program, changes due to the intervention, engagement with content, experience interacting with the supporter, experience of web-based communication, and termination of the supported period.

Conclusions: Patients who drop out of treatment can be distinguished in terms of their change in motivation: those who felt ready to leave treatment early and those who had negative reasons for dropping out. These 2 groups of participants have different treatment experiences, revealing the potential attributes and nonattributes of dropout. The reported between-group differences should be examined further to consider those attributes that are strongly descriptive of the experience and regarded less important than those that have become loosely affiliated.

(*JMIR Form Res* 2021;5(11):e26221) doi:[10.2196/26221](https://doi.org/10.2196/26221)

KEYWORDS

depression; anxiety; iCBT; dropout; internet interventions

Introduction

Background

The evidence base supporting internet-delivered cognitive behavioral therapy (iCBT) in the treatment of depression and anxiety is well established [1-5]. However, despite its apparent efficacy, treatment dropout continues to be reported throughout the literature [6]. Treatment dropout is most commonly defined as ending without completing treatment at a predefined cutoff point [7]. Some studies report that just more than half of patients complete a full course of iCBT [8], which is problematic in light of research associating lower completion rates with lower effect sizes [5]. At the same time, we know that it is not necessary for individuals to complete all the per-protocol treatment to benefit clinically [8]. Therefore, what remains is to try to understand dropouts distinctive from the binary classifications of adherence versus nonadherence to a given treatment protocol. Furthermore, considering that the goal of web-based psychological therapies is to provide an evidence-based and cost-effective treatment, help to reduce therapist time and waiting lists, compensate for lack of trained professionals, and alleviate the burden on mental health services of meeting demands [9], it is important to understand web-based treatment dropout to ensure maximal benefit for all involved.

To date, there has been a large body of quantitative research on dropout from web-based psychological therapies that has explored the associated variables to predict which individuals may be more at risk [6,7,10]. Male gender, lower educational levels, self-guided cognitive behavioral therapy (CBT) interventions, and depression with comorbid anxiety symptoms have each been found to significantly increase the risk of dropping out of web-based treatments, whereas the likelihood of dropping out significantly decreases with every additional 4 years of age [6,7]. Our current understanding is also informed by qualitative studies on adherence to internet-delivered psychological therapies [11-13]. These studies have found that the extent to which an individual adheres to a web-based treatment can be largely dependent on their assessment of the advantages and disadvantages of web-based delivery and how this meets their individual needs and preferences [11-13]. Adherers like the freedom and privacy provided by the web-based delivery of psychological interventions, have positive experiences with the content, trust the providers of the web-based program, are motivated to enroll in treatment, consider the web as a substitute for face-to-face support, and feel benefits from the intervention and their use is salient to their need for mental health interventions [11-13]. Their adherence does not appear to be attributed to a singular factor but to a collection of experiences, and their experiences are also not without negatives, citing difficulty with the language of the intervention, lacking confidence in the web-based delivery, a need for face-to-face support, and the potential for difficulties with technology. Although adherence research is a useful starting point for examining dropout, its application is limited insofar, as it cannot be assumed that the opposite will be true for dropout (ie, the disadvantages of web-based interventions explain dropout from them).

More recent research has focused exclusively on qualitatively analyzing individuals' experiences of web-based treatment dropout [14,15]. Johansson et al [15] reported that web-based treatment dropout is best understood in terms of an incompatible relationship between the perception of treatment and the patient's situation. A mismatch between any treatment feature and personal prerequisite results in the decision to nonadhere [15], for example, extensive and time-consuming content (treatment features), life factors such as commitments and availability (personal prerequisite), or the lack of face-to-face contact (treatment feature) and the personal preference for a need for face-to-face meetings. Reading and writing demands within the program (treatment feature) and individual capability at these tasks (personal prerequisites) is another example of a potential mismatch. Although a broad picture of web-based treatment dropout begins to emerge across the literature, with each study identifying different influential factors on dropout or suggesting a definitive factor at play, this lack of consensus may allude to the need for a different approach to its exploration.

Efforts to further understand web-based dropout must also take into consideration the conceptualization of the dropout and the implications of this [16]. Högdahl et al [16] summarized that web-based dropout seems to be conceptualized in terms of the number of modules completed rather than the effect of the treatment received on symptoms. This conceptualization may be problematic in light of research stating that treatment completion is not essential for clinical benefit [8]. Furthermore, other researchers argue that dropout is not necessarily a negative outcome or reflective of a wholly negative experience with a web-based intervention [17,18]. They call for patient discretion to be taken into consideration when evaluating and determining dropout status; patients may *drop out* of web-based interventions because they perceive their needs to have been met, and they no longer see the use of staying in treatment regardless of a predefined cutoff point [18,19].

Objective

The gaps in the web-based treatment dropout literature and the questioning of the current conceptualization suggests that dropout may be more nuanced, with individuals who meet dropout criteria in terms of a predefined number of sessions or modules having widely varied treatment experiences and motives for leaving treatment *prematurely* [17,18]. If this is the case, it is time for the investigation of dropout to move beyond predictors and individual reasons for dropout and look at the whole experience, incorporating the body of existing findings. It is important to ask *dropouts* about their experiences with treatment, exploring all potentially associated factors.

This study aims to conduct an in-depth exploration of the subjective experience of web-based treatment dropout by incorporating current literature on treatment dropout and adherence in both face-to-face and web-based contexts to create a robust semistructured interview. By interviewing and qualitatively analyzing individuals' experiences of dropout from an iCBT program in a routine care setting, it is hoped that a deeper insight into the experience of treatment dropout will be achieved.

Methods

Design

The study was a nested, semistructured qualitative interview study exploring clients' experiences of dropping out from an iCBT program for depression and anxiety [20]. It was part of a randomized controlled trial (RCT) investigating the effectiveness and cost-effectiveness of internet-delivered interventions for depression and anxiety in the United Kingdom's Improving Access to Psychological Therapies (IAPT) program [21]. The IAPT program is part of the National Health Service (NHS) designed to provide a stepped care approach for treating people with anxiety and depressive disorders. The results of this RCT showed that the intervention was effective at reducing symptoms of depression and anxiety compared with the waiting-list group, and these effects continued to improve over a 12-month follow-up. In addition, the RCT demonstrated that up to 60% of participants no longer met the criteria for a diagnosis of depression or anxiety at 3 months. With regard to the cost-effectiveness, the intervention was projected to be increasingly cost-effective across the 12-month follow-up horizon [21]. This study followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines for reporting qualitative research [22]. This study was approved by the NHS England research ethics committee (reference number 17/NW/0311).

Sample

The larger RCT included 361 individuals; of these 361 individuals, 66.8% (241/361) were randomized to the immediate treatment group and 33.2% (120/361) to the waiting-list control group. The design followed a 2:1 randomization procedure to reduce the likelihood of having many participants waiting for treatment after presenting to the IAPT service. All adult users of the Berkshire NHS Trust IAPT Talking Therapies step-2 services were eligible to participate. Clients were deemed

suitable for an internet intervention by their psychological well-being practitioner (PWP) based on their willingness to engage with an iCBT intervention, the presence of mild to moderate levels of anxiety or depression, no suicidal or self-harm risk, and having internet access. In line with the study protocol for the main RCT, a participant was considered to have dropped out of treatment if they received less than 6 web-based reviews from their supporter, as defined by the IAPT.

To identify eligible participants, the lead researcher (KL) manually went through each RCT participant's iCBT account history from the treatment group to verify the number of modules viewed, reviews received, and how responsive each of these participants was to their research contacts. Their level of responsiveness was determined by their history of answering calls from the RCT research team to complete the research measures. Eligibility criteria included (1) providing written informed consent, (2) completing fewer than 6 reviews with a supporter, and (3) completing a minimum of 1 module. The criterion of completing at least 1 module was necessary so that participants reporting on treatment dropout had some experience with each of the domains of investigation (see *Results*). A Microsoft Excel database was created listing 27 eligible participants for the qualitative interviews. Of the 27 eligible participants, 21 (78%) were invited to participate in the qualitative interviews via telephone at their 6-month or 9-month follow-up for the main RCT before 15 (56%) clients (of the 15 clients, 8/15, 53% women and 7/15, 47% were men) agreed to participate and were recruited. Purposive sampling [23] was used to recruit individuals for the semistructured interviews. Following the principles of purposive sampling, it was determined that after the 14th interview, there was a saturation of domains and categories; that is, no new information was being discovered [23]. This was confirmed by the results of the 15th interview. The mean age of participants was 33.5 (SD 9.1) years. The characteristics of the group are summarized in [Table 1](#).

Table 1. Characteristics of study participants^a.

Participant identifier	Gender	Age (years)	Mini-International Neuropsychiatric Interview diagnosis at baseline	iCBT ^b program	Modules completed, n (%)	Reviews received	Reported reason for change in motivation
P1	Female	24-26	Depression current or past	Space from Depression—8 modules (1 unlockable)	1 (13)	4	Negative reason (not in a receptive frame of mind, contextual obstacles, and iCBT not considered to be personally fitting)
P2	Female	50-53	Depression current or past, panic disorder, and GAD ^c	Space from Depression—8 modules (1 unlockable)	4 (50)	5	Negative reason (not in a receptive frame of mind, contextual obstacles, and iCBT not considered to be personally fitting)
P3	Female	34-36	No diagnosis	Space from Depression and Anxiety—10 modules (2 unlockable)	3 (30)	5	Felt ready to leave treatment early
P4	Female	24-26	Depression current or past	Space from Depression—8 modules (1 unlockable)	7 (88)	3	Felt ready to leave treatment early
P5	Male	30-33	GAD	Space from Depression and Anxiety—10 modules (2 unlockable)	5 (50)	3	Negative reason (iCBT not considered to be personally fitting)
P6	Male	37-39	Depression current or past	Space from Depression and Anxiety—10 modules (2 unlockable)	5 (50)	4	Negative reason (iCBT not considered to be personally fitting)
P7	Male	27-29	Depression current or past and GAD	Space from Depression and Anxiety—10 modules (2 unlockable)	1 (10)	3	Negative reason (not in a receptive frame of mind)
P8	Male	40-43	Depression current	Space from Depression and Anxiety—10 modules (2 unlockable)	7 (70)	2	Did not report
P9	Female	44-46	Panic disorder and GAD	Space from GAD—8 modules (1 unlockable)	4 (50)	2	Negative reason (not in a receptive frame of mind and iCBT not considered to be personally fitting)
P10	Male	44-46	Depression current or past, GAD, and SAD ^d	Space from Depression and Anxiety—10 modules (2 unlockable)	3 (30)	1	Negative reason (iCBT not considered to be personally fitting)
P11	Male	20-23	Depression past	Space from Depression and Anxiety—10 modules (2 unlockable)	4 (40)	4	Felt ready to leave treatment early
P12	Male	20-23	GAD	Space from Depression—8 modules	3 (38)	5	Negative reason (contextual obstacles, and iCBT not considered to be personally fitting)
P13	Female	37-39	No diagnosis	Space from Depression and Anxiety—10 modules (2 unlockable)	1 (10)	4	Did not report
P14	Female	34-36	GAD	Space from GAD—8 modules (1 unlockable)	3 (38)	5	Felt ready to leave treatment early
P15	Female	20-23	Depression current, panic disorder, GAD, and SAD	Space from GAD—8 modules (1 unlockable)	2 (25)	1	Felt ready to leave treatment early

^aParticipants have been allocated participant identifiers P1-P15 to protect their anonymity.

^biCBT: internet-delivered cognitive behavioral therapy.

^cGAD: generalized anxiety disorder.

^dSAD: social anxiety disorder.

Treatment

Space from Depression, *Space from Anxiety*, and *Space from Depression and Anxiety* are iCBT interventions for the treatment of depression and anxiety developed by SilverCloud Health with established efficacy [21,24]. These web-based programs comprise 5 core modules: Getting Started introduces CBT and the Thought Feeling Behavior (TFB) cycle, Understanding Feelings focuses on the *feelings* component of the TFB cycle, Spotting Thoughts focuses on the *thoughts* component of the TFB cycle, Challenging Thoughts focuses on taking action against negative and distorted thoughts, and Bringing it All Together prepares the user for coming to the end of the program [24]. *Space from Depression* has 2 additional modules: Boosting Behavior focuses on the inactivity and lack of motivation associated with depression, and Core Beliefs targets the underlying root of unhelpful thoughts that keep the cycle of depression going. *Space from Anxiety* has 2 additional modules: Facing Your Fears focuses on the role of avoidance in maintaining fears and anxiety, and Managing Worry focuses on recognizing real or hypothetical worries and identifying strategies to manage. All modules comprise cognitive and behavioral components, such as self-monitoring, thought recording, behavioral activation, and cognitive restructuring, along with incorporating relaxation exercises and personal stories from past users of the program to help guide clients on how to adapt the learned cognitive and behavioral strategies into their own lives [24].

The programs also use supporters that monitor patients' progress and provide asynchronous postsession feedback; this is referred to as a *review*. Reviews provide answers to patients' questions, encouragement and support on work completed and their progress, and signposts them to content. A dashboard interface gives supporters an overview of participants' level of engagement with the program. Clients in the study from which we recruited our participants were supported by a PWP from the Berkshire NHS Trust IAPT Talking Therapies service. PWPs

are graduate psychologists with further training in delivering low-intensity CBT-based interventions [25].

Measures

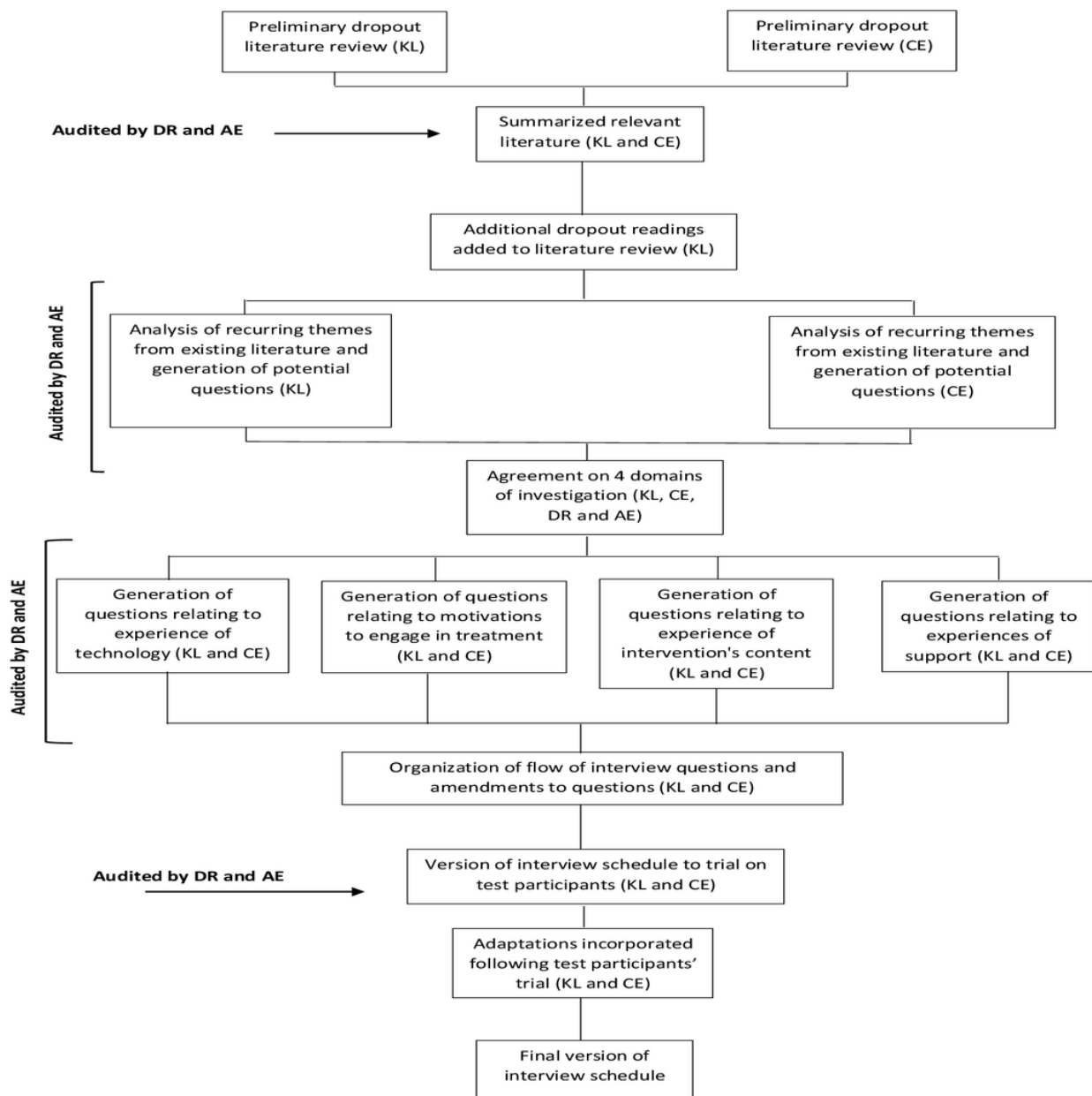
Mini-International Neuropsychiatric Interview 7.0

The Mini-International Neuropsychiatric Interview 7.0 is a short diagnostic structured interview based on both the Diagnostic and Statistical Manual of Mental Disorders and the International Classification of Diseases criteria. The interview and its administration by telephone have been well validated [26]. For its use during the main RCT, the interview schedule included modules A (major depressive episode), D (panic disorder), F (social anxiety disorder), and N (generalized anxiety disorder) to establish current depression and anxiety and specific anxiety presentations. The Mini-International Neuropsychiatric Interview 7.0 [26] diagnosis was used to assign participants to *Space from Depression*, *Space from Anxiety*, or *Space from Depression and Anxiety*, with the most suitable intervention being chosen for the participant based on symptomology.

Development of the Interview Schedule

KL and CE reviewed and analyzed the existing literature on treatment dropout to identify the recurring domains of investigation [7,10-15,27]. AE and DR audited this analysis, and 4 broad domains of investigation for treatment dropout were identified: experience of technology, motivations to engage in treatment, experience of intervention's content, and experiences of support. Questions were generated for each domain of investigation, balancing the greatest number of topics with the least number of questions. The interview was designed in line with the 4 main domains of investigation, and it was concluded that once these 4 domains were interviewed, there would be adequate information to address the research objectives. After discussion, questions were amended and selected for the interview schedule and organized within each domain before trialing the interview with a test participant. The interview schedule was refined once more before the final version was completed (Figure 1).

Figure 1. Stages of formation of interview schedule. Author initials are provided parenthetically.



Semistructured Interview

The interview ([Multimedia Appendix 1](#)) comprised 22 questions divided into 4 sections: 14% (3/22) questions in the Experience of Technology section (eg, “Did you welcome the intervention being online, considering that you do/don’t use much technology?”), 23% (5/22) questions in the Motivations to Engage in Treatment section (eg, “We note that you completed x sessions and x modules, what changed in this motivation?”), 41% (9/22) questions in Experience of Intervention’s Content section (eg, “Did you feel like the content in the programme was relevant to you?”), and 27% (6/22) questions in the Experiences of Support section (eg, “Everything that you do on the platform you have the option to share with your supporter, how did it feel to communicate in this way?”). This interview schedule provides a flexible framework for the interviews, with scripted prompts for the interviewer. The prompts were included

on a side panel of the interview schedule to ensure that the interviewer covered all domains of investigation, checking them off as they went to avoid repetition if a question had already been addressed in a different domain. Prompts also encouraged participants to adequately explore their subjective experiences of treatment dropout and to expand on them if their responses were lacking or they found it difficult to remember.

Characteristics of Interviewers, Researchers, and Auditors

This research project was led by a researcher with a background in e-Mental Health, who conducted the interviews and analysis (KL). The team of auditors was made up of a researcher undertaking counseling training (CE), a postdoctoral clinical researcher (AE), and a senior researcher (DR) who were all members of the e-Mental Health Research group at the Trinity College Dublin. In addition, a psychologist with a humanistic

orientation (LT) who emphasized clients' agency and interest in psychotherapy research was also a member of the auditor team.

Procedure

The interviews were conducted by 2 researchers via telephone and lasted between 27 minutes and 67 minutes, depending on the extent to which each participant explored their own experience of treatment dropout. The interviews were recorded and transcribed verbatim by a third-party transcription service. The web-based program was free to access, and participants received a £20 (US \$27.32) gift voucher for their participation in the interviews.

Data Analysis

Data were analyzed using descriptive and interpretive qualitative research methods [28] led by KL. The results were discussed and reflected upon with CE, LT, AE, and DR to ensure clarity and consensus on interpretations of the data and their meanings. The method of analysis followed clear steps.

First, the data were divided into discrete meaning units that captured the essence of what participants were trying to convey [29], and irrelevant digressions and repetitions were omitted. Meaning units were coded according to the order in which they occurred and to which participants they belonged. This process provided a clear audit path. All coding of the data was done manually, and Microsoft Excel was used to organize and store the data.

Second, meaning units were assigned to the domains of investigation headings (Experiences of Technology, Motivations to Engage in Treatment, Experiences of Intervention's Content, and Experiences of Support) to organize the data. The preliminary literature review that informed the creation of the semistructured interview schedule that was used for this study suggested domains of investigation; however, these were not finalized until after the data analysis.

Third, meaning units within the domains were grouped into categories based on their similar meanings. Some meaning units were included in more than 1 category, as they contained more than 1 relevant meaning (therefore, the categories are not mutually exclusive). For example, the meaning unit P2.29 stated the following:

[the supporter] kind of made suggestions...but I didn't feel [they] was imposing anything on me...[they] emailed something to me that wasn't on the platform[...] so I really felt they were taking their time to think of what I was going through.

The meaning unit was included in both categories—the category titled *Supporter offered understanding* and the category titled *Supporter tailored treatment to needs*. This process of categorization is subjective and interactive. The data are organized in a way that corresponds with the participants' meanings while also acknowledging the impact of existing theoretical knowledge [29], as outlined in the background of the interviewers, researchers, and auditors.

Fourth, strategies were used to maintain rigor and credibility. The first author (KL) divided the data into discrete meaning

units, and audits were performed at various intervals by the other authors to review this process. The process of organizing meaning units into domains and categories was conducted in several phases. The meaning units were first grouped into domains, and these choices were then discussed with fellow researchers who were experts in the literature, methodology, and iCBT, revising as necessary until agreement was reached. The same method was followed for categorization: the meaning units within each domain were grouped into categories and then presented to fellow researchers for comments and feedback. This process would be repeated until consensus was reached. Records were maintained for each step of the analysis. The feedback provided sometimes outlined a need for clarification of particular meaning units or their reallocation. Using this feedback sometimes resulted in the creation of new domains and categories or the removal of existing domains and categories.

Finally, during data analysis, 2 distinct participant groups emerged, characterized by reasons for the change in motivation to engage with the iCBT treatment: (1) those who felt ready to leave treatment early (5/15, 33%) and (2) those who had negative reasons for their change in motivation (8/15, 53%). Of the 15 participants, 2 (13%) did not report on the reason for their change in motivation to engage with treatment, and so they were excluded from the between-group comparison; however, they contributed to the formation of the overall domains and categories (Table S1 in [Multimedia Appendix 2](#)). To compare dropout experiences between the 2 groups, frequency labels were used, as outlined in the consensual qualitative research method [30]. We outline the representativeness of individual categories by considering them *general* if results apply to all cases (ie, 5/5 and 8/8 cases), *typical* if results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* if results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases) [30].

Results

A total of 10 domains capturing the areas of investigation of the subjective experiences of dropout from an iCBT intervention were formulated: relationship to technology, motivation to start, background knowledge and attitudes toward iCBT, change in motivation, use of the program, perceived changes because of the intervention, engagement with content, experience interacting with the supporter, experience of web-based communication, and termination of the supported period (Table S1 in [Multimedia Appendix 2](#)). Within each domain, there were positive and negative connotations for the participants' reports.

Change in Motivation

Participants who felt ready to leave treatment early (5/15, 33%) reported that they felt they had already obtained what they needed from the treatment without finishing the prescribed number of sessions:

I think it's just that point I sort of felt like I was getting better. I sort of got what I needed out of [the program]...I was feeling a bit better in my jowls and I didn't think I really needed it too much. [P15]

I got out of it what I needed and...the [supporter] I was speaking to gave me the option just to carry on logging on (my own)...I'm quite comfortable with logging on. [P14]

Participants who had negative reasons for their change in motivation (8/15, 53%) responded across 3 categories: not being in a receptive frame of mind, contextual obstacles, and considering iCBT not to be personally fitting:

But also and perhaps because I was just, my brain was just full up of loads of things going on I just wasn't in a receptive frame of mind. [P9]

I wasn't receptive enough to it at the time, but I do think in that frame of mind of feeling so low that you're kind of not...for months my brain didn't feel it was working very well. [P2]

Relationship to Technology

Relationship to technology corresponds to technology literacy, familiarity, and usability, both in general and specific to the iCBT program. Participants' reported relationships with their use of technology in general and with the technology itself were clustered into 10 categories that had both positive and negative connotations (Table 2). For the most part, both groups reported having positive relationships with technology: being familiar with technology, having a sense of anonymity and privacy on the web, and finding the platform easy to use. Negative relationships with technology were considered as a variant or not reported at all: difficulty figuring out how to use the web-based platform, spending too much time on the web, and having poor computer literacy.

Table 2. Participants' relationships to technology based on their reported reasons for their change in motivation.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Relationship with technology		
Being familiar with technology	General	General
Sense of privacy and anonymity on the web	General	General
Good memorability	General	Typical
Trusted the platform	Typical	Typical
Easy-to-use web-based platform	Typical	Typical
Spends too much time on the web	None	Variant
User dashboard not clear enough	None	Variant
Layout too structured	None	Variant
Difficulty figuring out how to use it	Variant	Variant
Poor computer literacy	Variant	None

^aOnly 13 participants (5/13, 38% felt ready to leave treatment early, and 8/13, 62% had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^bGeneral results apply to all cases (ie, 5/5 and 8/8), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8)

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered to be personally fitting.

Motivation to Start

This category referred to the reasons why participants sought mental health treatment in the first place. All participants reported on their motivations to seek treatment, and their responses were clustered into 2 categories (Table 3)—symptoms of psychological distress and stressful life events:

It was the most severe bout of depression that I've experienced. And it scared me, like I felt like I was having thoughts and reacting to things in a way that I couldn't control. [P4, felt ready]

So my husband had just left and I was panicking about like financially I didn't know what was gonna happen. [P2, negative reason]

Table 3. Participants' motivation to start treatment based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Motivation to start		
Symptoms of psychological distress	General	General
Stressful life events	Variant	Typical

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

Stressful life events as a motivation to start treatment was a typical category for participants who had negative reasons for their change in motivation but a variant category for participants who felt ready to leave treatment early, indicating that there were some differences between groups with regard to the motivation to start treatment.

take my own reflective time and think through my problems. [P5, negative reason]

I think I thought that [CBT] could change the way I think. [P12, negative reason]

There were between-group differences across the 3 categories (Table 4). A willingness to try the treatment, having no prior knowledge or awareness of CBT, and being skeptical of the treatment approach were typical to participants who felt ready to leave treatment early but variant to those who had negative reasons for their change in motivation:

Background Knowledge and Attitudes Toward iCBT

This category was characterized by what participants knew and believed about the iCBT program. Overall, the belief that iCBT could help was typical to both groups:

So when I started the sessions...I thought it would work really well for me because it would be [able to]

I didn't really know what CBT was [...] to a certain extent I didn't recognize that it would be so much about my thought processes and how it works. [P3, felt ready]

Table 4. Participants' background knowledge and attitudes toward iCBT^a based on their reported reasons for their change in motivation^b.

Domain and categories	Felt ready to leave treatment early (n=5) ^{b,c}	Negative reason for their change in motivation (n=8) ^{b,c,d}
Background knowledge and attitudes toward iCBT		
Belief that iCBT could help	Typical	Typical
Willingness to try it	Typical	Variant
Had an understanding of CBT ^e	Variant	Variant
Trusted provider of web-based treatment	Variant	Variant
No prior knowledge or awareness of CBT	Typical	Variant
Skeptical of treatment approach	Typical	Variant

^aiCBT: internet-delivered cognitive behavioral therapy.

^bGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^cOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^dReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting

^eCBT: cognitive behavioral therapy.

Use of the Program

This category was characterized by reports on how, why, and when participants used the program (Table 5). When use practices were compared between groups, productive and regular use practices were a general category for participants who felt ready to leave treatment early but a variant category for

participants who reported negative reasons for their change in motivation:

I set [a reminder] up for like every day at seven o'clock or something...When I'm sitting doing nothing it just gave me a little suggestion to go and do it, I guess. [P11, felt ready]

Table 5. Participants' use of the program based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Use of the program		
Could use it wherever and whenever needed	General	General
Productive and regular use	General	Variant
Using the program for own benefit	Typical	Variant
Could not prioritize time to use it	Typical	Typical
Using it out of a sense of obligation rather than for a positive outcome	Variant	Typical
Using it when feeling low	Variant	Typical
Kept forgetting about the program and appointments	Variant	Variant

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

It was typical for those who had negative reasons for their change in motivation to use it out of obligation or when feeling low:

It felt like obligation. It felt like a tick box exercise.
[P6, negative reason]

Perceived Changes Because of the Intervention

Participants' perceived changes because of the intervention, that is, new skills they acquired and changes to themselves and their everyday lives, were all positive (Table 6). Perceived symptom improvement was viewed as a general category for

those who felt ready to leave treatment early but a typical category for those who had negative reasons for their change in motivation:

When my dad did pass away because I was aware of all this stuff that I've learned [from the intervention]...And I purposefully the following week, on the exact same day, just to make sure that it [my OCD] wasn't there, I wore the exact same outfit. To push myself...to prove a point that it's got nothing to do with what I'm wearing, like it doesn't matter, it won't change it. [P3, felt ready]

Table 6. Participants' perceived changes because of the intervention based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Perceived changes because of the intervention		
Symptom improvement	General	Typical
Applying learned CBT ^d techniques in everyday life	Typical	Typical
Developed a knowledge of CBT treatment	Typical	Variant
Increased awareness or insight	Variant	Variant
Encouraged to get the help needed	None	Variant

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

^dCBT: cognitive behavioral therapy.

Conversely, being encouraged to get the help they needed was deemed a variant category for those who had negative reasons

for their change in motivation, whereas it was not reported by any participants who felt ready to leave treatment early:

I think it was definitely a benefit to kind of like dip my toes in and just get a feel for...cognitive behaviour therapy...it was definitely a good starting point for me. [P12, negative reason]

Engagement With Content

This category was characterized by reports of what participants liked and disliked about aspects of content within the program

(Table 7). There were some differences in reporting between the groups. Reflecting back on work completed being beneficial and writing about thoughts and feelings being therapeutic were typical to those who felt ready to leave treatment early:

[I reflected] sometimes, 'cos if I was having a really bad day and it wasn't as bad before, it made me feel a little bit better. [P14, felt ready]

Table 7. Participants' engagement with content based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Engagement with content		
Useful tools and exercises	Typical	General
Reflecting back on completed work was beneficial	Typical	Variant
Content relevant and relatable to concerns	Typical	Variant
Manageable workload	Variant	Variant
Reading and writing provided clarity	Variant	Variant
Writing about thoughts and feelings felt therapeutic	Typical	Variant
Felt supported by the program content	Typical	Variant
Information laid out clearly and concisely	Variant	Variant
Felt like too much work	Variant	Variant
Disliked reading and writing	Variant	Variant
Content was too generic at times	Variant	Variant
Did not like the personal stories	Variant	Variant
Content was boring	None	Variant
Content exacerbated symptoms	None	Variant
Reflecting of no benefit	None	Variant
Difficult to understand	None	Variant
Questionnaires felt pointless	None	Variant
Did not like the mood monitor	Variant	None
Content felt disconnected from one section to the next	None	Variant

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

However, these categories were deemed variant among those who had negative reasons for their change in motivation. Concerning negative experiences engaging with content, reports were low across both groups. Differences emerged with categories such as reflecting being of no benefit, content feeling disconnected between sections and difficult to understand, questionnaires feeling pointless, and finding the content boring or exacerbating to symptoms. There were variant categories among those who had negative reasons for their change in motivation but were not reported at all by those who felt ready to leave treatment early:

[the content] was a bit long winded to be honest with you. There was probably too much reading. So I probably skipped bits. [P10, negative reason]

Experience Interacting With the Supporter

This category relates to participants' comments on their relationship with their supporters and how they felt that interaction contributed to their overall treatment experience. Participants described these experiences across positive and negative dimensions (Table 8). Feeling supported and connected to their supporter was a general category for those who felt ready to leave treatment early but a variant category for those who had negative reasons for their change in motivation:

I recognise that I'm not looking someone in the face but it turns out to be the same to me because I still felt supported in everything that I did [...] there was

just someone there and that to me, was really good.
[P3, felt ready]

Table 8. Participants' experience interacting with supporters based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Experience interacting with supporter		
Felt supported by and connected to supporter	General	Variant
Supporter tailored treatment to needs	Typical	Typical
Supporter provided a good introduction and explanation of treatment	Variant	Typical
Felt able to speak freely	Typical	Variant
Supporter encouraged engagement	Typical	Variant
Benefitted from having a supporter	Typical	Variant
Supporter demonstrated a good level of expertise	Typical	Variant
Supporter discussed treatment goals	Variant	Variant
Supporter offered understanding	Variant	Variant
Support felt scripted and impersonal	None	Variant
Had no sense of connection with supporter	None	Variant
No feedback from supporter on work completed or messages sent	Variant	Variant
Supporter never discussed treatment goals and expectations	None	Variant
Lack of empathy and understanding from supporter	None	Variant
Lack of guidance from supporter	None	Variant
Felt like supporter did not care	None	Variant
Supporter never made contact	None	Variant
Did not feel comfortable talking with supporter	None	Variant

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4/5 and 5-7/8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

Furthermore, feeling able to speak freely with their supporter and the supporter demonstrating a good level of expertise was typical to those who felt ready to leave treatment early, whereas these categories were variant to those who had negative reasons for their change in motivation:

They do help you sort of really, really open up and you've got to remember, you know, they do this every single day. So, it was quite easy to open up in the first session. [P15, felt ready]

Having no connection with the supporter was a variant category for participants with negative reasons for their change in motivation but was not reported by those who felt ready to leave treatment early:

If I had felt a bit more that somebody was really listening and engaging [maybe we could have had a connection]. I just found it hard to build any sort of relationship. [P9, negative reason]

Although there was low reporting across the other negative categories, the same pattern applied between groups as with the lack of connection with the supporter.

Experience of Web-Based Communication

This category was characterized by participants' likes and dislikes on using a web-based medium to communicate with a supporter. Participants' reports relating to the medium of web-based communication were described across positive and negative categories, with large differences between groups (Table 9). Liking to communicate on the web with the supporter and finding it easier to open up on the web was typical to those who felt ready to leave treatment early compared with being variant categories for those who had negative reasons for their change in motivation:

I preferred [the online reviews] to be honest. And it was easy enough to do as well. [P14, felt ready]

Table 9. Participants' experience of web-based communication based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Experience of web-based communication		
Frequency of web-based communication worked well	Typical	Typical
Liked communicating web-based with supporter	Typical	Variant
Easier to open up on the web and feeling of disinhibition	Typical	Variant
Preference for face-to-face communication	None	Typical
Needed more contact with supporter	Variant	Variant
Communicating on the web was too formal and structured	None	Typical
Lack of instantaneous responding with supporter	None	Variant
Could not open up to a computer	None	Variant
Web-based communication felt too anonymous	None	Variant

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

Conversely, a preference for face-to-face interactions was a typical category for those who had negative reasons for their change in motivation but not reported at all by those who felt ready to leave treatment early:

I think looking back that maybe I should've had both, even though I was short on time, actually the (face-to-face) probably would've been better than maybe moving to [iCBT]. [P1, negative reason]

Termination of Supported Period

This category was characterized by participants' reports relating to how the supported period of the iCBT program was discontinued and how they felt about it (Table 10). It was typical for participants from both groups to feel able to go back to this treatment if they felt the need in the future:

If I explored a different route and it didn't work out then I was always welcome to rejoin SilverCloud, or rejoin Talking Therapies....So that is really positive. [P5, negative reason]

Being happy with how the supported period ended was a general category for those who felt ready to leave treatment early but

was not reported at all by those who had negative reasons for their change in motivation:

I think I got out of it what I needed...[my supporter] gave me the option just to carry on logging on (without support) and I'm quite comfortable with logging on. [P14, felt ready]

More negative reports relating to the termination of the supported period were variant categories for those who had negative reasons for their change in motivation, whereas they were not reported by those who felt ready to leave treatment early: support stopping unexpectedly and feeling abandoned and feeling relieved that support stopped because of it being such a negative experience:

[Support] stopped. I heard nothing, done nothing...I was shocked and disappointed. [P10, negative reason]

I haven't got time for this, you're not useful enough to me. Therefore, I'm not wanting to carry it on and give you my time because my time was too precious and as I say it just wasn't useful enough. [P9, negative reason]

Table 10. Participants' experience of termination of the supported period based on their reported reasons for their change in motivation^a.

Domain and categories	Felt ready to leave treatment early (n=5) ^{a,b}	Negative reason for their change in motivation (n=8) ^{a,b,c}
Termination of supported period		
Feels able to go back to treatment if needed	Typical	Typical
Happy with how support was terminated	General	None
Had a conversation with supporter about finishing treatment	Variant	Variant
No longer a priority, just let it go	Variant	Variant
Support stopped unexpectedly, felt abandoned	None	Variant
Felt relieved that support stopped as it was a negative experience	None	Variant

^aGeneral results apply to all cases (ie, 5/5 and 8/8 cases), *typical* results apply to at least half of the cases (ie, 3-4 of 5 and 5-7 of 8 cases), and *variant* results apply to fewer than half of the cases (ie, 1-2 of 5 and 1-4 of 8 cases).

^bOnly 13 participants (5/13, 38% participants felt ready to leave treatment early, and 8/13, 62% participants had negative reasons for their change in motivation) reported on the reasons for their change in motivation.

^cReported negative reasons for change in motivation to continue engaging with treatment are not being in a receptive frame of mind, contextual obstacles, and internet-delivered cognitive behavioral therapy not considered personally fitting.

Discussion

Principal Findings

This study qualitatively investigated dropout from iCBT interventions for depression and anxiety as part of routine mental health service delivery. It explored dropout across a continuum of 10 experiential domains. These domains are multiple and varied and demonstrate the conceptualization of treatment dropout as an experience not confined to one moment. Furthermore, this study establishes that when we conceptualize dropout in terms of the number of sessions completed, there are 2 distinct groups of participants: those with negative reasons for their change in motivation and those who feel ready to leave treatment early. However, the differences in treatment experiences observed between these groups point to a potential shift in how we think about treatment dropout.

This study has taken a deeper dive into dropout from iCBT treatment research. Previously, personal characteristics, individual capabilities, aspects of technology, intervention content, relationship with the supporter, motivation and treatment expectancies, and credibility have all been identified as reasons for treatment dropout [6,7,14,15]. This study examined dropout across the complete treatment experience and now recognizes it as a concept with related attributes and nonattributes, that is, treatment factors and experiences that describe it. The findings presented, although preliminary and requiring validation on a larger scale, advance what we know of dropout by suggesting that certain attributes of the treatment experience are strongly descriptive of dropout, whereas others may need to be reconsidered. To be descriptive of dropout, an attribute serves to describe the experiences of iCBT treatment leading up to dropout that were instrumental to the reason for dropout; therefore, a nonattribute is not pertinent to the decision to drop out.

Considering that participants described both their relationship with technology in general and to the iCBT program specifically as largely positive and that there were little to no differences in reporting between the 2 groups, it should be further considered whether technology is now a nonattribute of treatment dropout. To date, the literature has reported it to play an important role in dropout in terms of technology literacy, attitudes toward the technologization of health care, and technical difficulties [7,11-14,27]. Perhaps this is because after decades of development, design and technical flaws have been rectified [31,32], and technology and the internet have become both pervasive and accessible. These findings were echoed in the background knowledge and attitudes toward the iCBT programs domain, rendering it a nonattribute of treatment dropout. Individuals were mostly accepting of the use of technology in the delivery of their mental health care. Perhaps the routine care setting where iCBT is offered as a reliable treatment alternative may have acted as a buffer against the nonacceptance of and negative attitudes toward internet interventions [33-35].

Stressful life events before beginning iCBT treatment seem to be an attribute of dropout because proportionately more participants who had negative reasons for their change in motivation reported them than those who felt ready to leave treatment early. This is not surprising, as one of the main characteristics of this group is not continuing with treatment because of contextual obstacles such as work, relationships, and commitments. The literature has previously documented the influence of external factors on treatment dropout [8,10,15,36], stating that the demands they place on the individual will lead to dropout if viewed as an obstacle to their daily life [15]. Although these participants reported these life events as triggers for seeking and beginning treatment, they may have contributed to suboptimum conditions to continue treatment. The descriptive characteristics of stressful life events for dropouts have

implications for the clinical application of iCBT treatment. They may indicate a need for greater consideration to be given to suitability screening or increased pretherapy effort and tailoring (including content type and duration of treatment) of treatment to counteract them.

On the basis of the findings reported in relation to *perceived changes because of the intervention*, all participants perceived some benefit from the iCBT program, with the most commonly reported change being perceived symptom improvement, an attribute consistent with the delivery of an effective treatment in any format. This preliminary finding may align with the Waller and Gilbody [8] conclusion that treatment completion is not always necessary for clinical benefit and gives support to the Eysenbach [18] and Proudfoot [37] hypothesis that dropout is not necessarily reflective of a wholly negative experience. Furthermore, they suggest that perceived changes because of the intervention are attributes of treatment dropout. Although there may be reason to question the exclusivity of positive outcomes to treatment completion, these findings are tentative and largely based on patient discretion, and therefore, further research would be required to reach a conclusive decision regarding the relationship between dropout and positive outcomes. Interestingly, proportionately more participants with negative reasons for their change in motivation reported that the iCBT treatment encouraged them to receive the help they needed than those who felt ready to leave treatment early. This was considered a positive outcome as effective treatments work to leverage aspects of treatment, which prompts further help-seeking behaviors [38,39]. Further studies are required to explore this relationship between dropout and positive outcomes.

Negative interactions with a supporter and a lack of a connection characterize the dropout experience of those who had negative reasons for their change in motivation and who, according to the consensual qualitative research method categorizations [40], reported more than their counterparts across these categories. The differences observed between the 2 groups in terms of their experiences with supporters validate the emphasis placed on the patient-clinician relationship for treatment success [41,42] in the existing literature [7,10,11,13-15,27]. Although it is evident from previous research and the reports of those who felt ready to leave treatment early that this alliance can be established on the web [41,43], it may have been the case for participants who had negative reasons for their change in motivation that they could not overcome the altered dynamics of moving the relationship on the web [17,44]. From the findings of this study, it is clear that a poor-quality clinician-patient relationship facilitates dropout, rendering it an attribute.

There were also differences in reporting between groups in relation to the experience of web-based communication. A dislike for web-based communication and a preference for face-to-face treatment characterizes the dropout experiences of those who had negative reasons for their change in motivation. This finding is reflective of the pattern in reporting experiences with the supporter, and it would be interesting to investigate whether they are correlated. The role played by preferences in treatment dropout has been identified previously, concluding that despite the comparable efficacy of web-based communication with a supporter, an overwhelming number of

patients only prefer face-to-face interaction [43,45-47]. Considering the presence of such preferences in this study and previous research stating that if the support or communication type is not compatible with the patient's preferences or expectations, the patient may decide to drop out of treatment [15,48], a dislike for web-based communication and a preference for face-to-face treatment is a strongly descriptive of dropout, positing it as an attribute of treatment dropout. Further research should explore whether this dislike for web-based communication is a true dislike or whether it speaks more to the fact that, for these individuals, web-based communication is just not enough and should be provided in combination with face-to-face interaction.

The multiplicity and variance of the domains presented in this study expand our understanding of dropout. This nuanced portrayal was achieved through a robust methodology consisting of the development of a semistructured interview based on the existing literature pertaining to dropout and adherence in both face-to-face and web-based therapies and rigorous analysis using the descriptive-interpretive method [28]. It demonstrates the complexity of the dropout experience, calling into question the validity of conceptualizing treatment dropout in terms of compliance to modules or sessions needing completion, as some research does [16,49]. This has implications for both iCBT and clinical practice. As presented in the comparison between the 2 groups, some aspects of the treatment experience have become highly relevant to dropout: stressful life events before beginning treatment, using the iCBT program when feeling low or out of a sense of obligation, perceived changes because of the intervention, negative experiences with content, negative experiences with the supporter, a dislike for web-based communication, and a preference for face-to-face therapy, whereas others have diminished into the background or disappeared entirely: relationship to technology, background knowledge and attitudes toward iCBT, and termination of the supported period. It is apparent that all participants who dropped out benefitted somewhat; what needs to be understood now is who benefits the most, and these findings can help guide this future research. Furthermore, as the future of digital health care depends on the increased understanding of such phenomena so that psychological interventions can continue to increase in accessibility while increasing specificity for the patient, research exploring the complete experience is necessary.

Limitations

Although the ecological validity provided by the IAPT setting was a strength of this study, it may also have positively skewed participant reports, as their suitability for iCBT would have been assessed before beginning treatment. In addition, some individuals did not want to participate in the dropout interview. By not capturing the experiences of these individuals, the data presented may be positively biased to the intervention. The between-group differences that have been identified are based on qualitative data from a sample of 15 participants who dropped out in response to open-ended questions rather than closed-ended questions that would have investigated the presence or absence of an experience and so should be considered tentative. Future research into dropout should focus on identifying these 2 groups of participants on a larger scale and quantitatively investigate

their outcomes. As with any qualitative interview study, the potential roles played by social desirability, historical reporting, and researcher subjectivity should be taken into account. However, the results for all 15 participants were analyzed first, and it was determined only afterward that a second analysis with the participants divided into 2 groups would be useful to avoid any potential bias from the researcher.

Conclusions

The data presented from the qualitative interviews provide insight into the subjective experiences of participants who dropped out from an iCBT treatment for depression and anxiety in a routine care setting. In doing so, it moved beyond the current understanding of treatment dropout as a seemingly negative outcome attributable to a singular event and presents it as a phenomenon that must be considered experientially. The findings bring to light a more nuanced picture of treatment dropout when looked at through the perspective of varied domains that shed light on the experience. This suggests that participants who drop out can be distinguished in terms of their change in motivation: those who felt ready to leave treatment

early and those who had negative reasons for dropping out. In doing so, it facilitated a comparison of treatment experiences that revealed potential attributes (stressful life events before beginning treatment, using the iCBT program when feeling low or out of a sense of obligation, perceived changes because of the intervention, negative experiences with content, negative experiences with the supporter, a dislike for web-based communication, and a preference for face-to-face therapy) and nonattributes (relationship to technology, background knowledge and attitudes toward iCBT, and termination of the supported period) of dropout. To understand why individuals drop out, these between-group differences should be examined to consider those features that are strongly descriptive of the experience and regard those that have become loosely affiliated with less importance. The evidence presented in this study stipulates that there is a difference between what we label as a dropout and what should actually be considered a dropout. Further work, either quantitative or exploratory, is needed to comprehensively develop a typology of dropout participants and potentially reconceptualize the phenomenon in this rapidly changing digital health care setting.

Acknowledgments

The authors wish to thank the psychological well-being practitioners at Berkshire National Health Service Foundation Trust for providing the supported service for the internet-delivered cognitive behavioral therapy program. The authors thank the employees of the clinical and innovation team at SilverCloud for assisting in the development of the interview schedule. The authors also thank the patients who volunteered their time and shared their experiences to participate in our study. The study was funded by SilverCloud Health and the Irish Research Council. Employees at SilverCloud Health and Trinity College Dublin managed the data collection, data analysis, and writing of the manuscript. All authors approved the decision to submit for publication. SilverCloud is a commercial organization that sells its digital programs to commissioners within the National Health Service, who provide the service free to patients through the Improving Access to Psychological Therapies program.

Authors' Contributions

KL had full access to the data and was responsible for the integrity of the data and the accuracy of the data analysis. The study was conceptualized and designed by KL, CE, DR, and AE, and data acquisition, analysis, and interpretation was led by KL, CE, and LT. Theoretical discussion and conceptual interpretation of the findings were provided by DR, AE, and LT. KL drafted the manuscript. Critical revision of the paper was provided by LT, DR, AE, and CE. All authors reviewed and approved the final manuscript for submission. KL is the recipient of an Irish Research Council Scholarship under the Enterprise Partnership Postgraduate Scheme (EPSPG/2019/504).

Conflicts of Interest

KL is a past employee of SilverCloud Health, the developer of computerized psychological interventions for depression, anxiety, stress, and comorbid long-term conditions. CE, DR, and AE are current employees of SilverCloud Health. LT serves as a research consultant for SilverCloud Health.

Multimedia Appendix 1

Semistructured interview schedule exploring the experience of dropout from an internet-delivered cognitive behavioral therapy intervention.

[[PDF File \(Adobe PDF File\), 672 KB - formative_v5i11e26221_app1.pdf](#)]

Multimedia Appendix 2

Participants' experiences of treatment based on their reported reasons for their change in motivation.

[[DOCX File, 27 KB - formative_v5i11e26221_app2.docx](#)]

References

1. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009;38(4):196-205. [doi: [10.1080/16506070903318960](https://doi.org/10.1080/16506070903318960)] [Medline: [20183695](https://pubmed.ncbi.nlm.nih.gov/20183695/)]
2. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. *J Anxiety Disord* 2018 Apr;55:70-78 [FREE Full text] [doi: [10.1016/j.janxdis.2018.01.001](https://doi.org/10.1016/j.janxdis.2018.01.001)] [Medline: [29422409](https://pubmed.ncbi.nlm.nih.gov/29422409/)]
3. Richards D, Enrique A, Palacios J. Internet-delivered cognitive behaviour therapy. In: *Cognitive Behavioral Therapy and Clinical Applications*. London: IntechOpen; 2017.
4. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. *Clin Psychol Rev* 2012 Jun;32(4):329-342. [doi: [10.1016/j.cpr.2012.02.004](https://doi.org/10.1016/j.cpr.2012.02.004)] [Medline: [22466510](https://pubmed.ncbi.nlm.nih.gov/22466510/)]
5. Wright JH, Owen JJ, Richards D, Eells TD, Richardson T, Brown GK, et al. Computer-assisted cognitive-behavior therapy for depression: a systematic review and meta-analysis. *J Clin Psychiatry* 2019 Mar 19;80(2):18r12188 [FREE Full text] [doi: [10.4088/JCP.18r12188](https://doi.org/10.4088/JCP.18r12188)] [Medline: [30900849](https://pubmed.ncbi.nlm.nih.gov/30900849/)]
6. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. *Psychol Med* 2015 Oct;45(13):2717-2726. [doi: [10.1017/S0033291715000665](https://doi.org/10.1017/S0033291715000665)] [Medline: [25881626](https://pubmed.ncbi.nlm.nih.gov/25881626/)]
7. Melville KM, Casey LM, Kavanagh DJ. Dropout from internet-based treatment for psychological disorders. *Br J Clin Psychol* 2010 Nov;49(Pt 4):455-471. [doi: [10.1348/014466509X472138](https://doi.org/10.1348/014466509X472138)] [Medline: [19799804](https://pubmed.ncbi.nlm.nih.gov/19799804/)]
8. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. *Psychol Med* 2009 May;39(5):705-712. [doi: [10.1017/S0033291708004224](https://doi.org/10.1017/S0033291708004224)] [Medline: [18812006](https://pubmed.ncbi.nlm.nih.gov/18812006/)]
9. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. *J Med Internet Res* 2006 Jun 23;8(2):e10 [FREE Full text] [doi: [10.2196/jmir.8.2.e10](https://doi.org/10.2196/jmir.8.2.e10)] [Medline: [16867965](https://pubmed.ncbi.nlm.nih.gov/16867965/)]
10. Barrett MS, Chua W, Crits-Christoph P, Gibbons MB, Casiano D, Thompson D. Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy (Chic)* 2008 Jun 01;45(2):247-267 [FREE Full text] [doi: [10.1037/0033-3204.45.2.247](https://doi.org/10.1037/0033-3204.45.2.247)] [Medline: [19838318](https://pubmed.ncbi.nlm.nih.gov/19838318/)]
11. Beatty L, Binnion C, Kemp E, Koczwara B. A qualitative exploration of barriers and facilitatorsto adherence to an online self-help intervention for cancer-related distress. *Support Care Cancer* 2017 Aug;25(8):2539-2548. [doi: [10.1007/s00520-017-3663-2](https://doi.org/10.1007/s00520-017-3663-2)] [Medline: [28299458](https://pubmed.ncbi.nlm.nih.gov/28299458/)]
12. Todkill D, Powell J. Participant experiences of an internet-based intervention and randomised control trial: interview study. *BMC Public Health* 2013 Oct 28;13:1017 [FREE Full text] [doi: [10.1186/1471-2458-13-1017](https://doi.org/10.1186/1471-2458-13-1017)] [Medline: [24165325](https://pubmed.ncbi.nlm.nih.gov/24165325/)]
13. Holst A, Nejati S, Björkelund C, Eriksson MC, Hange D, Kivi M, et al. Patients' experiences of a computerised self-help program for treating depression - a qualitative study of Internet mediated cognitive behavioural therapy in primary care. *Scand J Prim Health Care* 2017 Mar;35(1):46-53 [FREE Full text] [doi: [10.1080/02813432.2017.1288813](https://doi.org/10.1080/02813432.2017.1288813)] [Medline: [28277055](https://pubmed.ncbi.nlm.nih.gov/28277055/)]
14. Fernández-Álvarez J, Díaz-García A, González-Robles A, Baños R, García-Palacios A, Botella C. Dropping out of a transdiagnostic online intervention: a qualitative analysis of client's experiences. *Internet Interv* 2017 Sep 22;10:29-38 [FREE Full text] [doi: [10.1016/j.invent.2017.09.001](https://doi.org/10.1016/j.invent.2017.09.001)] [Medline: [30135750](https://pubmed.ncbi.nlm.nih.gov/30135750/)]
15. Johansson O, Michel T, Andersson G, Paxling B. Experiences of non-adherence to Internet-delivered cognitive behavior therapy: a qualitative study. *Internet Interv* 2015 May;2(2):137-142. [doi: [10.1016/j.invent.2015.02.006](https://doi.org/10.1016/j.invent.2015.02.006)]
16. Högdahl L, Levallius J, Björck C, Norring C, Birgegård A. Personality predicts drop-out from therapist-guided internet-based cognitive behavioural therapy for eating disorders. Results from a randomized controlled trial. *Internet Interv* 2016 Jul 18;5:44-50 [FREE Full text] [doi: [10.1016/j.invent.2016.07.002](https://doi.org/10.1016/j.invent.2016.07.002)] [Medline: [30135806](https://pubmed.ncbi.nlm.nih.gov/30135806/)]
17. Proudfoot JG. Computer-based treatment for anxiety and depression: is it feasible? Is it effective? *Neurosci Biobehav Rev* 2004 May;28(3):353-363. [doi: [10.1016/j.neubiorev.2004.03.008](https://doi.org/10.1016/j.neubiorev.2004.03.008)] [Medline: [15225977](https://pubmed.ncbi.nlm.nih.gov/15225977/)]
18. 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/)]
19. Vandereycken W, Devidt K. Dropping out from a specialized inpatient treatment for eating disorders: the perception of patients and staff. *Eat Disord* 2010;18(2):140-147. [doi: [10.1080/10640260903585557](https://doi.org/10.1080/10640260903585557)] [Medline: [20390617](https://pubmed.ncbi.nlm.nih.gov/20390617/)]
20. Creswell J, Shope R, Plano CV. How interpretive qualitative research extends mixed methods research. *Res Sch* 2006;13(1):1-11.
21. Richards D, Enrique A, Eilert N, Franklin M, Palacios J, Duffy D, et al. A pragmatic randomized waitlist-controlled effectiveness and cost-effectiveness trial of digital interventions for depression and anxiety. *NPJ Digit Med* 2020 Jun 15;3:85 [FREE Full text] [doi: [10.1038/s41746-020-0293-8](https://doi.org/10.1038/s41746-020-0293-8)] [Medline: [32566763](https://pubmed.ncbi.nlm.nih.gov/32566763/)]
22. 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/)]
23. Silverman D. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*, 2nd edition. Thousand Oaks: SAGE Publications; 2001.

24. Richards D, Timulak L, O'Brien E, Hayes C, Vigano N, Sharry J, et al. A randomized controlled trial of an internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression. *Behav Res Ther* 2015 Dec;75:20-31. [doi: [10.1016/j.brat.2015.10.005](https://doi.org/10.1016/j.brat.2015.10.005)] [Medline: [26523885](https://pubmed.ncbi.nlm.nih.gov/26523885/)]
25. Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int Rev Psychiatry* 2011 Aug;23(4):318-327 [FREE Full text] [doi: [10.3109/09540261.2011.606803](https://doi.org/10.3109/09540261.2011.606803)] [Medline: [22026487](https://pubmed.ncbi.nlm.nih.gov/22026487/)]
26. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22-33;quiz 34. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
27. 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/)]
28. Elliott R, Timulak L. Descriptive and interpretive approaches to qualitative research. In: *A Handbook of Research Methods for Clinical and Health Psychology*. Oxford, UK: Oxford University Press; 2005.
29. McElvaney J, Timulak L. Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: an exploratory study. *Couns Psychother Res* 2013 Dec;13(4):1-8. [doi: [10.1080/14733145.2012.761258](https://doi.org/10.1080/14733145.2012.761258)]
30. Hill CE. *Consensual Qualitative Research: A Practical Resource for Investigating Social Science Phenomena*. Washington D.C: American Psychological Association; 2012.
31. Riley S, Veale D. The internet and its relevance to cognitive behavioural psychotherapists. *Behav Cogn Psychother* 1999 Jan 01;27(1):37-46. [doi: [10.1017/s1352465899271068](https://doi.org/10.1017/s1352465899271068)]
32. Selmi PM, Klein MH, Greist JH, Sorrell SP, Erdman HP. Computer-administered cognitive-behavioral therapy for depression. *Am J Psychiatry* 1990 Jan;147(1):51-56. [doi: [10.1176/ajp.147.1.51](https://doi.org/10.1176/ajp.147.1.51)] [Medline: [2403473](https://pubmed.ncbi.nlm.nih.gov/2403473/)]
33. Schröder J, Berger T, Meyer B, Lutz W, Hautzinger M, Späth C, et al. Attitudes towards internet interventions among psychotherapists and individuals with mild to moderate depression symptoms. *Cogn Ther Res* 2017 Apr 22;41(5):745-756. [doi: [10.1007/s10608-017-9850-0](https://doi.org/10.1007/s10608-017-9850-0)]
34. Díaz-García A, González-Robles A, Fernández-Álvarez J, García-Palacios A, Baños RM, Botella C. Efficacy of a transdiagnostic internet-based treatment for emotional disorders with a specific component to address positive affect: study protocol for a randomized controlled trial. *BMC Psychiatry* 2017 Apr 20;17(1):145 [FREE Full text] [doi: [10.1186/s12888-017-1297-z](https://doi.org/10.1186/s12888-017-1297-z)] [Medline: [28424068](https://pubmed.ncbi.nlm.nih.gov/28424068/)]
35. González-Robles A, García-Palacios A, Baños R, Riera A, Llorca G, Traver F, et al. Effectiveness of a transdiagnostic internet-based protocol for the treatment of emotional disorders versus treatment as usual in specialized care: study protocol for a randomized controlled trial. *Trials* 2015 Oct 31;16:488 [FREE Full text] [doi: [10.1186/s13063-015-1024-3](https://doi.org/10.1186/s13063-015-1024-3)] [Medline: [26519046](https://pubmed.ncbi.nlm.nih.gov/26519046/)]
36. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009 Apr 24;11(2):e13 [FREE Full text] [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
37. Proudfoot J, Ryden C, Everitt B, Shapiro DA, Goldberg D, Mann A, et al. Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry* 2004 Jul;185:46-54. [doi: [10.1192/bjp.185.1.46](https://doi.org/10.1192/bjp.185.1.46)] [Medline: [15231555](https://pubmed.ncbi.nlm.nih.gov/15231555/)]
38. Wampold B, Imel Z. *The Great Psychotherapy Debate The Evidence for What Makes Psychotherapy Work*. England: Taylor & Francis; 2015.
39. Kazdin AE. Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol* 2007;3:1-27. [doi: [10.1146/annurev.clinpsy.3.022806.091432](https://doi.org/10.1146/annurev.clinpsy.3.022806.091432)] [Medline: [17716046](https://pubmed.ncbi.nlm.nih.gov/17716046/)]
40. Hill C, Thompson B, Williams E. A guide to conducting consensual qualitative research. *Couns Psychol* 1997 Oct 1;25(4):517-572 [FREE Full text] [doi: [10.1177/0011000097254001](https://doi.org/10.1177/0011000097254001)]
41. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. *Psychother Theory Res Pract* 1979;16(3):252-260. [doi: [10.1037/h0085885](https://doi.org/10.1037/h0085885)]
42. Bordin ES. Theory and research on the therapeutic working alliance: new directions. In: *The Working Alliance: Theory, Research, and Practice*. Hoboken, New Jersey: John Wiley & Sons; 1994.
43. Berger T. The therapeutic alliance in internet interventions: a narrative review and suggestions for future research. *Psychother Res* 2017 Sep;27(5):511-524. [doi: [10.1080/10503307.2015.1119908](https://doi.org/10.1080/10503307.2015.1119908)] [Medline: [26732852](https://pubmed.ncbi.nlm.nih.gov/26732852/)]
44. Sucala M, Schnur JB, Brackman EH, Constantino MJ, Montgomery GH. Clinicians' attitudes toward therapeutic alliance in E-therapy. *J Gen Psychol* 2013;140(4):282-293 [FREE Full text] [doi: [10.1080/00221309.2013.830590](https://doi.org/10.1080/00221309.2013.830590)] [Medline: [24837821](https://pubmed.ncbi.nlm.nih.gov/24837821/)]
45. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. *J Clin Psychol* 2010 Apr;66(4):394-409 [FREE Full text] [doi: [10.1002/jclp.20659](https://doi.org/10.1002/jclp.20659)] [Medline: [20127795](https://pubmed.ncbi.nlm.nih.gov/20127795/)]

46. Wallin EE, Mattsson S, Olsson EM. The preference for internet-based psychological interventions by individuals without past or current use of mental health treatment delivered online: a survey study with mixed-methods analysis. *JMIR Ment Health* 2016 Jun 14;3(2):e25 [FREE Full text] [doi: [10.2196/mental.5324](https://doi.org/10.2196/mental.5324)] [Medline: [27302200](https://pubmed.ncbi.nlm.nih.gov/27302200/)]
47. Schröder J, Sautier L, Kriston L, Berger T, Meyer B, Späth C, et al. Development of a questionnaire measuring Attitudes towards Psychological Online Interventions-the APOI. *J Affect Disord* 2015 Nov 15;187:136-141. [doi: [10.1016/j.jad.2015.08.044](https://doi.org/10.1016/j.jad.2015.08.044)] [Medline: [26331687](https://pubmed.ncbi.nlm.nih.gov/26331687/)]
48. Rogers EM. *Diffusion of Innovations*, 5th edition. New York: Simon & Schuster; 2003.
49. Hynan DJ. Client reasons and experiences in treatment that influence termination of psychotherapy. *J Clin Psychol* 1990 Nov;46(6):891-895. [doi: [10.1002/1097-4679\(199011\)46:6<891::aid-jclp2270460631>3.0.co;2-8](https://doi.org/10.1002/1097-4679(199011)46:6<891::aid-jclp2270460631>3.0.co;2-8)] [Medline: [2286687](https://pubmed.ncbi.nlm.nih.gov/2286687/)]

Abbreviations

CBT: cognitive behavioral therapy
COREQ: Consolidated Criteria for Reporting Qualitative Research
IAPT: Improving Access to Psychological Therapies
iCBT: internet-delivered cognitive behavioral therapy
NHS: National Health Service
PWP: psychological well-being practitioner
RCT: randomized controlled trial
TFB: thought feeling behavior

Edited by G Eysenbach; submitted 03.12.20; peer-reviewed by J Coumans, A Parks, N Forand; comments to author 25.01.21; revised version received 16.03.21; accepted 01.08.21; published 12.11.21.

Please cite as:

Lawler K, Earley C, Timulak L, Enrique A, Richards D

Dropout From an Internet-Delivered Cognitive Behavioral Therapy Intervention for Adults With Depression and Anxiety: Qualitative Study

JMIR Form Res 2021;5(11):e26221

URL: <https://formative.jmir.org/2021/11/e26221>

doi: [10.2196/26221](https://doi.org/10.2196/26221)

PMID: [34766909](https://pubmed.ncbi.nlm.nih.gov/34766909/)

©Kate Lawler, Caroline Earley, Ladislav Timulak, Angel Enrique, Derek Richards. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.11.2021. 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 Online Patient Satisfaction Index for Patients With Low Back Pain: Development, Reliability, and Validation Study

Tamana Afzali¹, MSc; Henrik Hein Lauridsen², PhD; Janus Laust Thomsen¹, PhD; Jan Hartvigsen^{2,3}, PhD; Martin Bach Jensen¹, PhD; Allan Riis^{1,4}, PhD

¹Research Unit for General Practice in Aalborg, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

²Departments of Sports Science and Clinical Biomechanics, Centre for Muscle and Joint Health, University of Southern Denmark, Odense, Denmark

³Nordic Institute of Chiropractic and Clinical Biomechanics, Odense, Denmark

⁴Department of Physiotherapy, University College Northern Denmark, Aalborg, Denmark

Corresponding Author:

Allan Riis, PhD

Research Unit for General Practice in Aalborg

Department of Clinical Medicine

Aalborg University

Fyrkildevej 7

Aalborg, 9220

Denmark

Phone: 45 20823660

Email: ariis@dcm.aau.dk

Abstract

Background: Low back pain is highly prevalent, and most often, a specific causative factor cannot be identified. Therefore, for most patients, their low back pain is labeled as nonspecific. Patient education and information are recommended for all these patients. The internet is an accessible source of medical information on low back pain. Approximately 50% of patients with low back pain search the internet for health and medical advice. Patient satisfaction with education and information is important in relation to patients' levels of inclination to use web-based information and their trust in the information they find. Although patients who are satisfied with the information they retrieve use the internet as a supplementary source of information, dissatisfied patients tend to avoid using the internet. Consumers' loyalty to a product is often applied to evaluate their satisfaction. Consumers have been shown to be good ambassadors for a service when they are willing to recommend the service to a friend or colleague. When consumers are willing to recommend a service to a friend or colleague, they are also likely to be future users of the service. To the best of our knowledge, no multi-item instrument exists to specifically evaluate satisfaction with information delivered on the web for people with low back pain.

Objective: This study aims to report on the development, reliability testing, and construct validity testing of the Online Patient Satisfaction Index to measure patients' satisfaction with web-based information for low back pain.

Methods: This is a cross-sectional validation study of the Online Patient Satisfaction Index. The index was developed with experts and assessed for face validity. It was subsequently administered to 150 adults with nonspecific low back pain. Of these, 46% (70/150) were randomly assigned to participate in a reliability test using an intraclass correlation coefficient of agreement. Construct validity was evaluated by hypothesis testing based on a web app (MyBack) and Wikipedia on low back pain.

Results: The index includes 8 items. The median score (range 0-24) based on the MyBack website was 20 (IQR 18-22), and the median score for Wikipedia was 12 (IQR 8-15). The entire score range was used. Overall, 53 participants completed a retest, of which 39 (74%) were stable in their satisfaction with the home page and were included in the analysis for reliability. Intraclass correlation coefficient of agreement was estimated to be 0.82 (95% CI 0.68-0.90). Two hypothesized correlations for construct validity were confirmed through an analysis using complete data.

Conclusions: The index had good face validity, excellent reliability, and good construct validity and can be used to measure satisfaction with the provision of web-based information regarding nonspecific low back pain among people willing to access the internet to obtain health information.

Trial Registration: ClinicalTrials.gov NCT03449004; <https://clinicaltrials.gov/ct2/show/NCT03449004>

KEYWORDS

data accuracy; patient satisfaction; rehabilitation; low back pain; internet-based intervention; mobile phone

Introduction

Background

Low back pain (LBP) is highly prevalent and is the most frequent reason for patients to consult general practice in Denmark [1,2]. LBP affects men and women of all ages [3] and is rarely caused by one specific factor [4,5]. Therefore, for most patients, their LBP is labeled as nonspecific, that is, a nociceptive source is not well established, and causes are multifactorial [6,7]. Patient education and information are generally recommended for people seeking care for nonspecific LBP [8]. However, delivering evidence-based information can be time-consuming and cumbersome, which can be a challenge during the available consultation time in general practice [9].

The internet is an accessible source of medical information for patients and it offers a range of information provided by a variety of sources. It has been reported that approximately 50% of patients search the internet for health and medical advice [10], and evidence suggests that this is increasing [11]. The advancement of new technologies offers more opportunities for delivering patient information on private computers, tablets, and smartphones, and web-based information can be considered an inexpensive solution to extend the treatment in general practice [12]. Therefore, future optimization of web-based information delivery has the potential to increase the delivery of evidence-based information about LBP, which may, in turn, lead to better patient outcomes [13,14].

Patient satisfaction is important in the use of web-based information and the degree to which patients rely on information from the internet [15]. Although patients who are satisfied with the information they retrieve use the internet as a supplementary source of information, dissatisfied patients tend to avoid using the internet [11]. Consumers' loyalty to a product is often applied to evaluate their satisfaction [16]. Consumers have been shown to be good ambassadors for a service or product when they are willing to recommend the service to a friend or colleague [16,17]. When consumers are willing to recommend a service to a friend or colleague, they are also likely to be future users of the service [16,17].

Objectives

To the best of our knowledge, no multi-item instrument to specifically evaluate satisfaction with information delivered on the web for people with LBP exists. This study aims to report on the development and validation of the Online Patient Satisfaction Index (OPSI), a self-reported measure to evaluate patients' satisfaction with web-based information for LBP.

Methods

Overview

This study was registered at ClinicalTrials.gov (ID: NCT03449004). The study follows the Consensus-based Standards for the Selection of Health Measurement Instruments Taxonomy [18]. The *Methods* section consists of 2 subsections: development of the OPSI and validation of the OPSI.

Development of the OPSI

Previous Work

A qualitative interview study had previously identified patients' preferences for the content, design, and functionality of a web application with evidence-based information and advice for patients with LBP consulting general practice [19]. This study identified a set of important domains to address web-based information and advice for patients with LBP in Denmark and highlighted the importance of the following domains: design, readability, customization, credibility, usability, and coping [19]. On the basis of these findings, 8 specific items related to these domains were identified as important for patient satisfaction with web-based information for LBP. Design and readability were represented with 2 items; the other domains were presented with 1 item each.

Development Process

A total of 8 items were combined into the first version of the index. The content of the 8 items in the first version came from an interview study [19], after which the authors made a draft version where experts (not members of the author group) and 10 patients provided input. Thereafter, the reliability and validity were tested. All items initially had response options ranging from 0 to 10, where 0 indicated strongly dissatisfied and 10 indicated strongly satisfied. The first version of the index was then tested for face validity by discussing the wording of the items with 7 experts. The experts were personally invited among colleagues but were outside the author group (1 academic and 1 researcher experienced in written communication, 2 researchers with expertise in musculoskeletal disorders, and 3 researchers with expertise within the development of questionnaires). This process was carried out through 2 rounds, where the first draft was discussed among the experts and subsequently modified. After round 1, the questions were reformulated to reduce jargon and for better wording of the items, and the order of the items was rearranged to create a better flow of the index. This revised index was then discussed with the same experts until a consensus on the final version was reached. Importantly, the 0 to 10 response rate scale about satisfaction was found to be difficult to use, and therefore, the response scale was changed from the numerical rating scale to a categorical scale about satisfaction with 4 response options: *Very Much*, *Quite a bit*, *A little*, and *Not at all* (Figure 1).

Figure 1. The Online Patient Satisfaction Index, translated to English from Danish.

Here is a questionnaire about your satisfaction with the website. Please tick a box for every question.

If you find it hard to select an answer, we recommend that you choose the answer, which fits you best.

	Very Much	Quite a bit	A little	Not at all
1. Do you like the website's appearance?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Is it easy to find your way around the website?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Do you understand the content?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Is the language suitable for you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Is the content relevant for you?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Do you trust the website?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Did you find what you were looking for?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Is the website a good tool to help you with your back problem?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Conceptual Framework

The OPSI is based on a formative model in which the construct (satisfaction) is the result of patients' experiences with different aspects of satisfaction. For example, the items relating to design, credibility, and readability can all have an impact on patients' satisfaction with the web-based information, whereas higher satisfaction with a home page does not necessarily lead to patients finding it more customized to their needs [20].

Face Validity

The OPSI was pilot tested for face validity on 10 respondents with nonspecific LBP from the Sano Centre and 10 respondents with LBP recruited from social media. The Sano Centre is a training and rehabilitation center for people with a high degree of musculoskeletal pain or disability. First, they were asked to fill in baseline characteristics on paper for age, sex, pain duration (>12 weeks), pain intensity, curiosity to find new knowledge (0-10), and frequency of internet searching for health-related information (monthly or more). With 1 researcher (AR or TA) present, respondents were asked to search for information on an existing website (The Patient Handbook) [21]. The publicly available Patient Handbook has previously been found to be trustworthy and a preferred site among Danes searching for

information about LBP [19]. The author group was not involved in developing the design or choosing the content of the Patient Handbook. After assessing the Patient Handbook for 10 minutes, respondents were asked to complete the OPSI and were encouraged to comment openly on the process and content. Their thoughts and comments were noted on paper by the researcher. This was done to optimize the content validity of the items by reducing ambiguity, avoiding double-barreled questions, reducing jargon terms, reducing the length of the items, checking the existence of irrelevant items, and patients were asked if there was a lack of any items related to patient satisfaction with web-based information. We specifically asked about the feasibility of the questions, their understanding of the items, and the reasons for their choice of response options. Respondents' thoughts and suggestions about the index were discussed between TA and AR, and the index was revised and ready for validation among a larger population of respondents with nonspecific LBP. Figure 1 shows the English version of the OPSI, which has been forward-backward translated from Danish using the method suggested by Beaton et al [22] with modifications to stage 4. In stage 4, 1 native English-speaking researcher in expertise in musculoskeletal disorders achieved consensus with 2 native Danes holding a master's degree in

English. Although stages 1 to 3 were conducted at personal meetings, stage 4 was conducted on the web with TA as a facilitator. The Danish version is shown in [Figure 2](#).

Figure 2. The Online Patient Satisfaction Index in Danish. Each item contributes 0 to 3 points, giving a total Online Patient Satisfaction Index score between 0 and 24.

**Nedenunder finder du spørgsmål angående din tilfredshed med hjemmesiden.
Du bedes sætte ét kryds ved hvert spørgsmål. Hvis du har svært ved at vælge dit svar, opfordrer vi til at du vælger det svar, som passer bedst.**

	Meget	Noget	Lidt	Slet ikke
1. Kan du lide hjemmesidens udseende?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Er det nemt at finde rundt på hjemmesiden?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Forstår du indholdet?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Er sproget målrettet til dig?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Er indholdet relevant for dig?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Stoler du på hjemmesiden?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Fandt du det, du ledte efter?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Er hjemmesiden et godt værktøj til at hjælpe dig med dine rygproblemer?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Construct Validity and Reliability of the OPSI

Data from the validation study were gathered using paper versions of the OPSI, and TA entered the data in REDCap (Research Electronic Data Capture, Vanderbilt University) [23].

Participants

Sample sizes for validation studies are recommended to contain more than 100 participants and at least 50 participants per subgroup [20]. Consequently, the required sample size was set at 150 participants. The sample population for the development of the OPSI was recruited from Sano Centre in Aarhus, Denmark, and via social media (Facebook, LinkedIn, and Twitter) to obtain a case-mix. The patients had a 4-week stay at the center, and the training and rehabilitation course was patient-centered and tailored to the patients' needs. Patients were offered specialized therapy, and most received an iPad as part of their training program [24]. Hence, most people were familiar with the internet and electronic devices to manage their pain. Respondents from social media were also expected to be familiar with the use of the internet and electronic devices. Consequently, the sample population was expected to be heterogeneous regarding their levels of pain and functional disability.

Respondents from Sano and social media were eligible for inclusion if they had nonspecific LBP (with or without leg pain) of any pain intensity during the previous year and were older than 18 years. Exclusion criteria were as follows: no internet access, pregnancy, inability to speak Danish as their native language, diagnosis of spinal stenosis, or signs of a serious underlying disease (signs of fracture, cauda equina syndrome, malignancy, osteoporosis, or spondyloarthritis).

Procedures

The study was registered by the Danish Data Protection Agency (J.nr. 2017-41-5222). Ethics approval was not required following Danish law. Respondents received verbal project information from TA or AR, and informed consent was signed by the respondents and the assessor. Testing was performed at the Sano Centre, public libraries, or in respondents' homes with either TA or AR present. Initially, respondents filled in baseline questions on paper and were encouraged to navigate and search for a new home page, MyBack, for 10 minutes. The MyBack home page is in Danish and contains information about LBP to guide patients with self-management [25]. The content and design were developed by researchers with systematic input from patients and general practitioners [19,26,27]. Respondents

then filled in paper versions of the OPSI to measure their satisfaction with MyBack together with other questions about satisfaction and their functional disability. If any item was left blank, the assessor encouraged respondents to choose the response that was most suitable for them. Thereafter, respondents navigated Wikipedia with information about LBP for 10 minutes with minimum help from the assessor. We assumed that most respondents would be more satisfied with MyBack than with information about LBP on Wikipedia. As the amount of information regarding LBP is limited in the Danish version of Wikipedia, we considered it difficult to read, not addressed to a particular group of people, and the sources are unknown and might, therefore, not be considered trustworthy by respondents. Thereafter, respondents filled out the OPSI for the Wikipedia page.

Among the respondents, 70 were randomized to be invited to participate in the retest after a minimum of 7 days and a maximum of 4 weeks. Randomization and allocation numbers were provided by a researcher who was not involved in this study. A publicly available home page was used to generate 150 numbers with *yes* or *no* and sent the document with the allocation numbers to the assessors [28]. Respondents randomized to *yes* were invited to participate in a retest. Respondents who agreed to take part in the retest searched and navigated the MyBack home page a second time after a minimum of 7 days with the same assessor present (TA or AR). After filling out paper versions of the OPSI, respondents were encouraged to respond to 1 question about the stability of their satisfaction: Do you think your satisfaction with the home page has changed since last time? (Answer options: *yes/no/don't know*). Their replies were discussed with the assessor to validate their responses and if responding *no* or *don't know*, they were considered stable and included in the retest analysis.

Measurement Tools

A web app can be considered a product, and we assume that users are satisfied when they are more likely to recommend a web app to a friend or colleague. *The Ultimate Question* was used as the primary outcome measurement to compare measurement properties with the OPSI. *The Ultimate Question* is often applied to measure costumers' satisfaction with products or services using a single question: How likely are you to recommend the website to others? The question can be answered using a 0 to 10 scale, where 0 is the least likely to recommend and 10 is the most likely to recommend [16,17]. Respondents replying 10 or 9 are considered promoters and are likely to buy or use the product or service again, respondents replying 8 or 7 are considered passive, and respondents replying 6-0 are considered detractors [16,17]. Subtracting the percentage of detractors from the percentage of promoters yields the net promoter score (NPS).

For this study, we dichotomized the scores by defining the responses 10 or 9 as *promoters* of the web-based information and, therefore, considered satisfied, whereas we defined responses from 8 to 0 as *nonpromoters* and thereby not satisfied [16,17]. A Global Rating Scale (GRS) of satisfaction with the web app was applied (*What is your overall satisfaction with the*

website?) with response options ranging from 10 (very satisfied) to 0 (not at all satisfied).

The Danish version of the Roland-Morris Disability Questionnaire (RMDQ) was used to measure LBP related function [29]. The Danish version consists of 23 items, and the sum score ranges from 0 (no disability) to 23 (maximum disability) [29].

Statistical Evaluation

Reliability was assessed by studying the difference between the OPSI at baseline and OPSI after a minimum of 7 days using a 2-way mixed-effect analysis of variance model with interaction for absolute agreement as the intraclass correlation coefficient of agreement ($ICC_{\text{agreement}}$) [30]. Where an $ICC_{\text{agreement}} > 0.75$ can be interpreted as excellent, 0.4 to 0.75 indicate fair to poor, and values < 0.4 indicate poor reliability [31]. Measurement error was assessed using the limits of agreement proposed by Bland and Altman [32]. The smallest detectable change ($SDC_{\text{consistency}}$) was calculated as follows:

$$SDC_{\text{consistency}} = 1.96 \times SD_{\text{difference}}$$

where $SD_{\text{difference}}$ is the SD of the difference between the test and retest. This equals the limits of agreement without systematic errors [33]. Construct validity was evaluated by hypothesis testing of the size and direction of correlations between the OPSI score for MyBack and the NPS, GRS, RMDQ, and OPSI score for Wikipedia about LBP using Spearman rank correlation coefficient. CIs were estimated by bootstrapping with 5000 replications. Correlations between 0.3 and 0.5 were considered weak, and correlations > 0.5 were considered strong [34].

Hypothesis 1

We hypothesized that being categorized as *promoters* (scoring 9 or 10 on the NPS) would be positively and strongly (> 0.5) correlated with higher OPSI scores (convergent validity).

Hypothesis 2

We hypothesized that respondent scores from 9 to 10 on the GRS would be positively and strongly (> 0.5) correlated with higher OPSI scores (convergent validity).

Hypothesis 3

We hypothesized that RMDQ scores would be positively and weakly (< 0.3) correlated with higher OPSI scores (discriminant validity).

Hypothesis 4

We hypothesized that the OPSI score for MyBack and the OPSI score based on the Wikipedia website would be negatively and weakly (-0.3 to 0) correlated with higher OPSI scores for OPSI based on MyBack (discriminant validity).

All analyses were conducted using Stata 16.0 (Stata Corp).

Results

A total of 150 participants were recruited between March 6, 2018, and May 10, 2019. The mean age of the participants was

48.7 (SD 12.9) years, and 67.3% (101/150) were women. Most (146/150, 97.3%) had experienced pain for >12 weeks with an average score of 8 (range, 0-10) for having an interest in finding new information on the internet (Table 1).

Table 1. Characteristics of the study population (N=150).

Baseline characteristics	Patients
Age (years), mean (SD)	48.7 (12.9)
Female, n (%)	101 (67.3)
Education level, bachelor's degree or more, n (%)	55 (36.7)
Employment status, n (%)	
Working full-time or part-time	66 (44)
On sick leave or leave of absence	39 (26)
Unemployed	8 (5.3)
Retired	37 (24.7)
Pain duration >12 weeks, n (%)	146 (97.3)
Pain intensity, 0-10, mean (SD)	5.4 (2.1)
RMDQ ^a score, mean (SD)	11.8 (5.1)
Contact with GP ^b about LBP ^c during the past 1 year, n (%)	146 (97.3)
Curious about finding new information, 0-10, median (IQR)	8 (7-9)
Use of internet about health, n (%)	
Daily	9 (6)
Weekly	48 (32)
Less than weekly	10 (6.7)
Monthly	39 (26)
Less than monthly	44 (29.3)
OPSI ^d home page, median (IQR)	20 (18-22) ^e
OPSI Wikipedia, median (IQR)	12 (8-15) ^f
OPSI home page retest, median (IQR)	20 (17-22) ^g

^aRMDQ: Roland-Morris Disability Questionnaire.

^bGP: general practitioner.

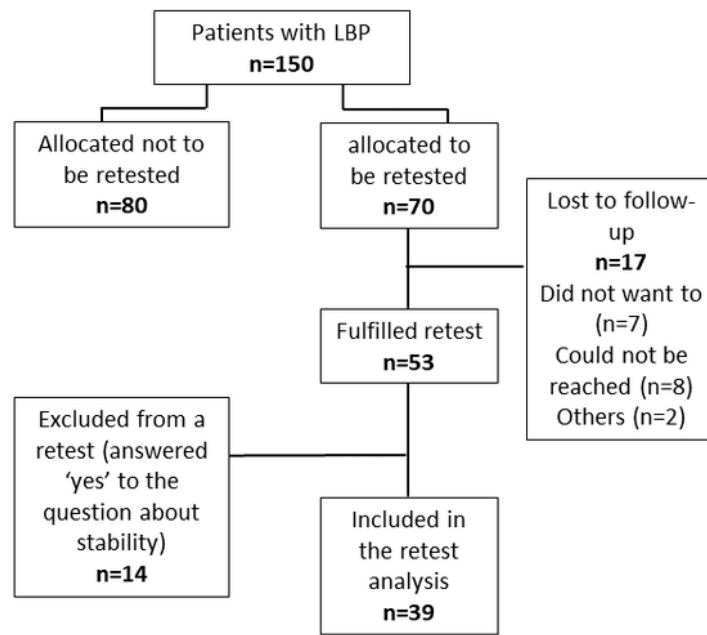
^cLBP: low back pain.

^dOPSI: Online Patient Satisfaction Index.

^{e,f,g}Distribution of the OPSI scores were nonnormal, which were identified on histograms and q-norm visualizations.

A total of 70 randomly chosen participants were invited to the retest, 53 accepted whereas 39 answered *no* or *don't know* to the question about the stability of their satisfaction. These 39 were considered stable and were included in a retest analysis to evaluate the stability of the OPSI (Figure 3).

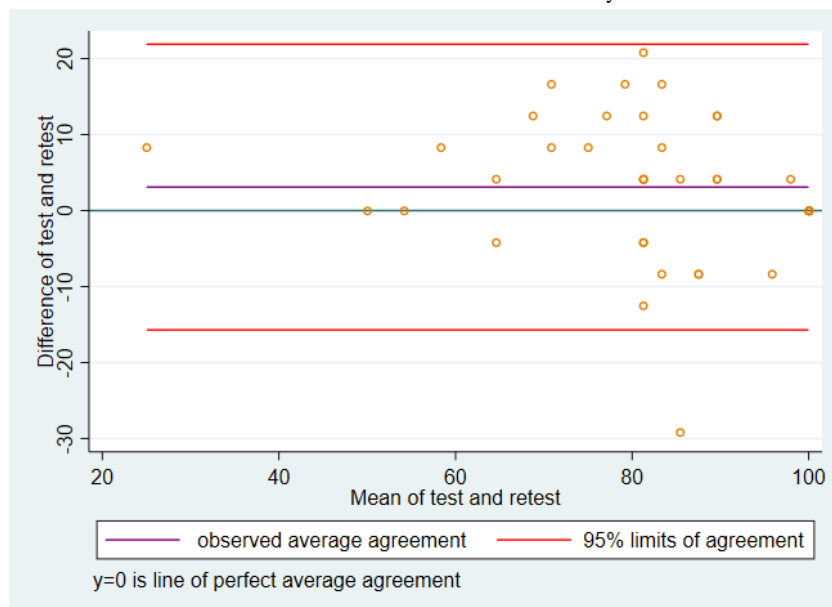
Figure 3. Flowchart of respondents. A total of 150 respondents were included in the validity analyses, and 39 were included in the reliability analysis. LBP: low back pain.



For the MyBack website, the OPSI score ranged from 724, and for the Wikipedia website, the entire OPSI score (0-24) range was used. The mean OPSI score of promoters for the MyBack website was 21.68 (95% CI 21.14-22.22), and the mean OPSI score of the nonpromoters was 18 (95% CI 17.25-18.75). The response rate was 100% for both websites. ICC_{agreement} was estimated at 0.82 (95% CI 0.68-0.90). The SDC_{consistency} was estimated at 4.71.

Limits of agreement were estimated to be -4.11 to 5.13 with a mean difference of 0.509 (95% CI -0.127 to 1.146). The mean difference is close to zero with the CI overlapping zero; hence, there is a negligible systematic difference between the baseline measurement and the retest measurement. In addition, the difference between baseline and follow-up measurements did not seem to depend on the level of satisfaction (Figure 4).

Figure 4. Comparison of test-scores and retest-scores. Online Patient Satisfaction Index for MyBack at the initial test and retested after 1 week.



Two hypotheses were confirmed (Table 2).

Table 2. Hypothesis testing of construct validity.

Characteristics	Hypotheses	Correlations ^a	
		Expected, <i>r</i>	Observed ^b , <i>r</i> (95% CI)
NPS ^c	Hypothesis A ^d	>0.5	0.58 (0.47 to 0.69)
GRS ^e	Hypothesis A	>0.5	0.70 (0.60 to 0.80)
RMDQ ^f	Hypothesis B ^g	<0.3	-0.14 (-0.30 to 0.02)
OPSI _{wikipedia} ^h	Hypothesis C ⁱ	-0.3 to 0	-0.04 (-0.20 to 0.13)

^aSpearman rank correlation coefficient.

^bCorrelations from confirmed hypotheses.

^cNPS: net promoter score.

^dScales are expected to measure the same construct. The correlation was expected to be positive and strong (>0.5).

^eGRS: Global Rating Scale. ^fRMDQ: Roland-Morris Disability Questionnaire.

^gScales are not expected to measure the same construct. The correlation was expected to be positive and weak (<0.3).

^hOPSI: Online Patient Satisfaction Index.

ⁱScales are expected to measure the same construct. The correlation is expected to be negative and weak (-0.3 to 0).

Discussion

Principal Findings

The OPSI is easy to apply as each of the 8 items is scored from 0 to 3, resulting in an overall score between 0 and 24 points. The index was found to have good face validity, excellent reliability, and good construct validity among our sample of participants with long-standing nonspecific LBP.

Recommendations to the OPSI Users

We recommend the use of proportional recalculation to convert the index score to a 0 to 100 scale, as it accounts for items with missing scores [35]. The OPSI was found to have excellent reliability in measuring satisfaction at 1 time point. Thus, it is suitable to measure satisfaction at 1 time point for 1 person or to compare satisfaction between groups. However, we did not test for responsiveness among participants experiencing a change in satisfaction over time; consequently, we cannot recommend using the OPSI to measure changes over time. Whether this can be recommended in the future, needs to be supported by the evaluation of responsiveness.

Limitations

The item *Do you trust the website?* is expected to capture both trust in the content and trust in security and handling of data. During development and face validity, splitting this item into two was not mentioned by participants. This may be explained by the high level of trust in health care authorities handling data in Denmark, but we do not know if this is the case. In other cultures, trust in the content and trust in the handling of data by the provider may be considered as 2 different issues and thereby require 2 items to be properly captured.

The development and evaluation of the OPSI was based on a formative model, which is 1 of the 2 conceptual frameworks, the other being a reflective model. A reflective model assumes interrelatedness between items and thereby item correlations. However, a formative model does not assume item correlations, and this is a limitation, as common statistical methods to

describe the relationships between items and the construct were not applicable in this study [20].

The Sano Centre receives a new cohort of patients every 4 weeks, most of whom have long-standing LBP, so it is convenient and easy to ask patients to participate in the study because they stay at the center. For participants volunteering over social media, the assessors had to make an appointment to meet at a convenient location, which was not always easy and straightforward. Consequently, this led to an unequal distribution of respondents between the 2 sites, with (131/150, 87.3%) from the Sano Centre and (19/150, 12.7%) from social media. This is a limitation of the study as patients from the Sano Centre were expected to have more severe symptoms, require extended information, and thus score a larger difference between the 2 home pages. In contrast, participants recruited from social media may be more frequent users of Wikipedia for other information seeking and thereby more satisfied with the shorter wiki format. The size and direction of the Spearman coefficient changed for patients recruited from social media. However, this could be due to the small number of patients recruited from social media.

The time between the test and retest in the analysis of reliability was between 1 and 2 weeks, with a maximum of 4 weeks, which might have overestimated the reliability of OPSI. The short time duration may have influenced the participants' responses as some may recall their previous baseline response and, therefore, repeat the answer. However, people with back pain often experience changes in symptoms, and the short period between tests can be considered a strength when collecting data from participants with possible fluctuations in symptoms within a few weeks [36]. We applied a stability question to determine whether the participants were considered stable. It is a strength that TA or AR discussed with the participants to ensure that participants understood the question on change in satisfaction regarding MyBack and to ensure that participants with a change were excluded from the retest. We used the construct of *satisfaction* as a reflection of whether the included respondents in the reliability study were stable, and this is in fact treating

the items as a reflective model. However, the alternatives would either be to ask the respondents if they had changed each item or not to ask them at all. The first alternative would cause analytic problems determining who were stable and who were not, as some respondents probably would have changed on 1 or perhaps 2 items but not the rest. Setting a cutoff point on an acceptable number of stable items would be arbitrary and probably misleading, and we also question the feasibility of doing it this way. The second alternative of not asking about *stability* is, in our opinion, unacceptable, as the potential to introduce a bias is high. We therefore opted for the solution of implementing a *global change question* as we believe this is the best of the 3 options introducing least bias regarding choosing stable respondents. However, assuming *don't know* as stable is potentially a limitation.

The use of NPS as a comparator for the construct validity of OPSI to measure satisfaction might be a weakness. The NPS uses a proxy to assess the customer's overall satisfaction with a service or product [17]. NPS is based on only 1 question with a reply option from 0 to 10, which is categorized into 3 groups: promoters (10-9), passive (8-7), and detractors (6-0) [17]. When estimating satisfaction, the middle group (passive) was excluded from the analysis [17]. This is, in our opinion, a weakness that can limit the usability of the NPS as a comparator for construct validity. We also applied the GRS as a comparator for construct validity using a response scale ranging from 0 to 10. The GRS may be inflated due to recall bias [37] and motivational effects [38] in longitudinal studies. Furthermore, transition scores seldom show an ideal pattern of association between baseline and follow-up measures [37]. However, we used baseline GRS to compare with a baseline score for the OPSI, and using a transition score to compare baseline scores has previously been found feasible [37].

Comparison With Prior Work

Although we have found no previous studies evaluating satisfaction with web-based information on LBP, other studies have reported satisfaction with web-based delivered health care information and educational information. Hence, a recent study of satisfaction with a gamified medical course among medical students in Thailand found a mean satisfaction of 9.02 (SD 1.11) out of 10 [39]. Equal high satisfaction was recorded in a study about the satisfaction of SMS text containing educational material for patients undergoing prostate biopsy in the United States. This study found a mean satisfaction of 4.5 (SD 0.9) out of 5 [40]. Furthermore, a high level of satisfaction was found in a study of satisfaction with telephone support in patients with type 2 diabetes [41,42]. They recorded a mean of 4/5 on all 21 items related to satisfaction [42]. These previous studies indicate a problem with ceiling effects when measuring satisfaction with health care interventions. The median score of OPSI used on a home page, which we considered good, was 20 [18,22] out of a maximum of 24 points. This indicates that satisfied patients tended to reply at the high end. Furthermore, this is in line with the development and use of the NPS, where only customers scoring 9 or 10 on a scale from 0 to 10 are considered satisfied

with a product or a service to a degree where customers are likely to buy a product again or reuse the service [16].

The reliability test indicated that a change of 4.71 points is necessary to preclude measurement error. If performing proportional rescaling of OPSI to a scale from 0 to 100, a change of approximately 20 points is necessary to preclude measurement error. In another study, applying a 0 to 100 score, the minimal detectable change was found to be 32.8 points on the group level when measuring the individualized quality of life in patients with LBP [43]. In a reliability and responsiveness study evaluating the minimal detectable change in LBP disability, questionnaires found lower or similar minimal detectable changes of 8.6 (RMDQ), 15 (Oswestry Disability Questionnaire), and 22 (the 36-item Short Form health surveys physical functioning scale) [44]. Consequently, compared with other measures applied to patients with LBP, OPSI has a similar good reproducibility.

Unanswered Questions and Future Research

The OPSI was evaluated in a population with severe and disabling LBP. Consequently, the index will benefit from further validation in other populations with less severe symptoms. In addition, the OPSI can be tested as a more generic tool to assess satisfaction without including dissatisfaction. King et al [45] constructed a survey with 22 items about the usability of phone apps to support physical activity and rated it on a 6-point Likert-type scale [45]. This scale was later adapted into a 21-item questionnaire rated for agreement or disagreement on a 5-point Likert scale [46]. A questionnaire developed more recently to measure user satisfaction with mobile health (mHealth) apps also applied a Likert scale—from *strongly disagree* to *strongly agree* [47]. However, this questionnaire did not meet the criteria for unidimensionality [47]. The OPSI is not challenged by the same potential limitation, as only satisfaction is measured and dissatisfaction is not.

We did not collect specific data regarding health literacy; however, either volunteering to participate on the internet or participating in a 4-week course at Sano indicates a health-interest in seeking health information among participants. This interest is supported by participants' baseline characteristics regarding the use of the internet to search for health information. Engaging all people, particularly those with low health literacy, in assessing health information can be a challenge [48]. Nevertheless, future research is needed to evaluate whether the OPSI can also be used by people with low health literacy. The index was developed and formally evaluated in Danish; however, it was translated into English using established guidelines [22]. The OPSI is potentially applicable in all Western countries, but future studies need to evaluate validity and reliability among other cultures and in other languages.

Conclusions

The OPSI showed good face validity, excellent reliability, and good construct validity and can be used when measuring satisfaction with the provision of information regarding LBP among people willing to access the internet for health information.

Acknowledgments

The Novo Nordisk Foundation (NNF17OC0024422) and the Danish Rheumatism Association (R132-A3476-B750) funded this study. They have no role in the study design, collection of data, management, analysis, interpretation of data, writing of the report, or the decision to submit the report for publication.

Conflicts of Interest

None declared.

References

1. Kent PM, Keating JL. The epidemiology of low back pain in primary care. *Chiropr Osteopat* 2005 Jul 26;13:13 [FREE Full text] [doi: [10.1186/1746-1340-13-13](https://doi.org/10.1186/1746-1340-13-13)] [Medline: [16045795](https://pubmed.ncbi.nlm.nih.gov/16045795/)]
2. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. *Lancet* 2018 Jun 09;391(10137):2356-2367. [doi: [10.1016/S0140-6736\(18\)30480-X](https://doi.org/10.1016/S0140-6736(18)30480-X)] [Medline: [29573870](https://pubmed.ncbi.nlm.nih.gov/29573870/)]
3. Hoy D, Brooks P, Blyth F, Buchbinder R. The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol* 2010 Dec;24(6):769-781. [doi: [10.1016/j.berh.2010.10.002](https://doi.org/10.1016/j.berh.2010.10.002)] [Medline: [21665125](https://pubmed.ncbi.nlm.nih.gov/21665125/)]
4. Bardin LD, King P, Maher CG. Diagnostic triage for low back pain: a practical approach for primary care. *Med J Aust* 2017 Apr 03;206(6):268-273. [doi: [10.5694/mja16.00828](https://doi.org/10.5694/mja16.00828)] [Medline: [28359011](https://pubmed.ncbi.nlm.nih.gov/28359011/)]
5. Lemeunier N, Leboeuf-Yde C, Gagey O. The natural course of low back pain: a systematic critical literature review. *Chiropr Man Therap* 2012 Oct 17;20(1):33 [FREE Full text] [doi: [10.1186/2045-709X-20-33](https://doi.org/10.1186/2045-709X-20-33)] [Medline: [23075327](https://pubmed.ncbi.nlm.nih.gov/23075327/)]
6. Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. *Lancet* 2012 Feb 04;379(9814):482-491. [doi: [10.1016/S0140-6736\(11\)60610-7](https://doi.org/10.1016/S0140-6736(11)60610-7)] [Medline: [21982256](https://pubmed.ncbi.nlm.nih.gov/21982256/)]
7. Vlaeyen JW, Maher CG, Wiech K, Van Zundert J, Meloto CB, Diatchenko L, et al. Low back pain. *Nat Rev Dis Primers* 2018 Dec 13;4(1):52. [doi: [10.1038/s41572-018-0052-1](https://doi.org/10.1038/s41572-018-0052-1)] [Medline: [30546064](https://pubmed.ncbi.nlm.nih.gov/30546064/)]
8. Low back pain and sciatica in over 16s: assessment and management. National Institute for Health and Care Excellence. 2016. URL: <https://www.nice.org.uk/guidance/ng59> [accessed 2021-10-15]
9. Bishop FL, Dima AL, Ngui J, Little P, Moss-Morris R, Foster NE, et al. "Lovely Pie in the Sky Plans": a qualitative study of clinicians' perspectives on guidelines for managing low back pain in primary care in England. *Spine (Phila Pa 1976)* 2015 Dec;40(23):1842-1850. [doi: [10.1097/BRS.0000000000001215](https://doi.org/10.1097/BRS.0000000000001215)] [Medline: [26571064](https://pubmed.ncbi.nlm.nih.gov/26571064/)]
10. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients' use of the internet for medical information. *J Gen Intern Med* 2002 Mar;17(3):180-185 [FREE Full text] [doi: [10.1046/j.1525-1497.2002.10603.x](https://doi.org/10.1046/j.1525-1497.2002.10603.x)] [Medline: [11929503](https://pubmed.ncbi.nlm.nih.gov/11929503/)]
11. Tustin N. The role of patient satisfaction in online health information seeking. *J Health Commun* 2010 Jan;15(1):3-17. [doi: [10.1080/10810730903465491](https://doi.org/10.1080/10810730903465491)] [Medline: [20390974](https://pubmed.ncbi.nlm.nih.gov/20390974/)]
12. Tailored on-line information and advice to patients with low back pain in general practice: a protocol for a semi-structured interview study (Part of the ADVIN Back Trial). Aalborg University. 2016. URL: https://vbn.aau.dk/ws/portalfiles/portal/241172247/Interview_study_2909_2016.pdf [accessed 2021-10-15]
13. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, Lancet Low Back Pain Series Working Group. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet* 2018 Jun 09;391(10137):2368-2383. [doi: [10.1016/S0140-6736\(18\)30489-6](https://doi.org/10.1016/S0140-6736(18)30489-6)] [Medline: [29573872](https://pubmed.ncbi.nlm.nih.gov/29573872/)]
14. Buchbinder R, van Tulder M, Öberg B, Costa LM, Woolf A, Schoene M, Lancet Low Back Pain Series Working Group. Low back pain: a call for action. *Lancet* 2018 Jun 09;391(10137):2384-2388. [doi: [10.1016/S0140-6736\(18\)30488-4](https://doi.org/10.1016/S0140-6736(18)30488-4)] [Medline: [29573871](https://pubmed.ncbi.nlm.nih.gov/29573871/)]
15. Al-Abri R, Al-Balushi A. Patient satisfaction survey as a tool towards quality improvement. *Oman Med J* 2014 Jan;29(1):3-7 [FREE Full text] [doi: [10.5001/omj.2014.02](https://doi.org/10.5001/omj.2014.02)] [Medline: [24501659](https://pubmed.ncbi.nlm.nih.gov/24501659/)]
16. Reichheld F, Markey R. *The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-driven World*. Massachusetts: Harvard Business Press; 2011.
17. Reichheld FF. The one number you need to grow. *Harv Bus Rev* 2003 Dec;81(12):46-54, 124. [Medline: [14712543](https://pubmed.ncbi.nlm.nih.gov/14712543/)]
18. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010 Jul;63(7):737-745. [doi: [10.1016/j.jclinepi.2010.02.006](https://doi.org/10.1016/j.jclinepi.2010.02.006)] [Medline: [20494804](https://pubmed.ncbi.nlm.nih.gov/20494804/)]
19. 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/)]
20. Henrica CW, Lidwine BM, Dirk LKC. *Measurement in Medicine: A Practical Guide (Practical Guides to Biostatistics and Epidemiology)*. Cambridge, UK: Cambridge University Press; 2011.
21. The patient handbook. Sundhed. URL: <https://www.sundhed.dk/borger/patienthaandbogen> [accessed 2021-10-15]

22. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000 Dec 15;25(24):3186-3191. [doi: [10.1097/00007632-200012150-00014](https://doi.org/10.1097/00007632-200012150-00014)] [Medline: [11124735](https://pubmed.ncbi.nlm.nih.gov/11124735/)]
23. 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/)]
24. Velkommen til Sano. Sano Danish Rheumatism Association. URL: <https://sanocenter.dk> [accessed 2021-10-15]
25. Min ryg: information om rygsmerter. Aalborg University. 2017. URL: <https://vbn.aau.dk/en/publications/min-ryg-information-om-rygsmerter> [accessed 2021-10-15]
26. Hjelmager DM, Vinther LD, Poulsen SH, Petersen LS, Jensen MB, Riis A. Requirements for implementing online information material for patients with low back pain in general practice: an interview study. *Scand J Prim Health Care* 2019 Mar;37(1):60-68 [FREE Full text] [doi: [10.1080/02813432.2019.1569413](https://doi.org/10.1080/02813432.2019.1569413)] [Medline: [30700191](https://pubmed.ncbi.nlm.nih.gov/30700191/)]
27. Riis A, Hartvigsen J, Rathleff MS, Afzali T, Jensen MB. Comparing satisfaction with a participatory driven web-application and a standard website for patients with low back pain: a study protocol for a randomised controlled trial (part of the ADVIN Back Trial). *Trials* 2018 Jul 25;19(1):399 [FREE Full text] [doi: [10.1186/s13063-018-2795-0](https://doi.org/10.1186/s13063-018-2795-0)] [Medline: [30045749](https://pubmed.ncbi.nlm.nih.gov/30045749/)]
28. Random.org. URL: <https://www.random.org/> [accessed 2021-10-28]
29. Albert HB, Jensen A, Dahl D, Rasmussen MN. [Criteria validation of the Roland Morris questionnaire. A Danish translation of the international scale for the assessment of functional level in patients with low back pain and sciatica]. *Ugeskr Laeger* 2003 Apr 28;165(18):1875-1880. [Medline: [12772398](https://pubmed.ncbi.nlm.nih.gov/12772398/)]
30. Qin S, Nelson L, McLeod L, Eremenco S, Coons SJ. Assessing test-retest reliability of patient-reported outcome measures using intraclass correlation coefficients: recommendations for selecting and documenting the analytical formula. *Qual Life Res* 2019 Apr;28(4):1029-1033 [FREE Full text] [doi: [10.1007/s11136-018-2076-0](https://doi.org/10.1007/s11136-018-2076-0)] [Medline: [30547346](https://pubmed.ncbi.nlm.nih.gov/30547346/)]
31. Fleiss J. *The Design and Analysis of Clinical Experiments*. USA: John Wiley & Sons; 1999.
32. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999 Jun;8(2):135-160. [doi: [10.1177/096228029900800204](https://doi.org/10.1177/096228029900800204)] [Medline: [10501650](https://pubmed.ncbi.nlm.nih.gov/10501650/)]
33. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006 Oct;59(10):1033-1039. [doi: [10.1016/j.jclinepi.2005.10.015](https://doi.org/10.1016/j.jclinepi.2005.10.015)] [Medline: [16980142](https://pubmed.ncbi.nlm.nih.gov/16980142/)]
34. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York: Routledge; 1988.
35. Kent P, Lauridsen HH. Managing missing scores on the Roland Morris Disability Questionnaire. *Spine (Phila Pa 1976)* 2011 Oct 15;36(22):1878-1884. [doi: [10.1097/BRS.0b013e3181ffe53f](https://doi.org/10.1097/BRS.0b013e3181ffe53f)] [Medline: [21304432](https://pubmed.ncbi.nlm.nih.gov/21304432/)]
36. Kongsted A, Kent P, Axen I, Downie AS, Dunn KM. What have we learned from ten years of trajectory research in low back pain? *BMC Musculoskelet Disord* 2016 May 21;17:220 [FREE Full text] [doi: [10.1186/s12891-016-1071-2](https://doi.org/10.1186/s12891-016-1071-2)] [Medline: [27209166](https://pubmed.ncbi.nlm.nih.gov/27209166/)]
37. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. *J Clin Epidemiol* 2002 Sep;55(9):900-908. [doi: [10.1016/s0895-4356\(02\)00435-3](https://doi.org/10.1016/s0895-4356(02)00435-3)] [Medline: [12393078](https://pubmed.ncbi.nlm.nih.gov/12393078/)]
38. Aseltine RH, Carlson KJ, Fowler FJ, Barry MJ. Comparing prospective and retrospective measures of treatment outcomes. *Med Care* 1995 Apr;33(4 Suppl):AS67-AS76. [Medline: [7536868](https://pubmed.ncbi.nlm.nih.gov/7536868/)]
39. Phungoen P, Promto S, Chanthawatthanarak S, Maneepong S, Apiratwarakul K, Kotruchin P, et al. Precourse preparation using a serious smartphone game on advanced life support knowledge and skills: randomized controlled trial. *J Med Internet Res* 2020 Mar 09;22(3):e16987 [FREE Full text] [doi: [10.2196/16987](https://doi.org/10.2196/16987)] [Medline: [32149711](https://pubmed.ncbi.nlm.nih.gov/32149711/)]
40. Balakrishnan AS, Nguyen HG, Shinohara K, Au Yeung R, Carroll PR, Odisho AY. A mobile health intervention for prostate biopsy patients reduces appointment cancellations: cohort study. *J Med Internet Res* 2019 Jun 02;21(6):e14094 [FREE Full text] [doi: [10.2196/14094](https://doi.org/10.2196/14094)] [Medline: [31199294](https://pubmed.ncbi.nlm.nih.gov/31199294/)]
41. McGloin H, O'Connell D, Glacken M, Mc Sharry P, Healy D, Winters-O'Donnell L, et al. Patient empowerment using electronic telemonitoring with telephone support in the transition to insulin therapy in adults with type 2 diabetes: observational, pre-post, mixed methods study. *J Med Internet Res* 2020 May 14;22(5):e16161 [FREE Full text] [doi: [10.2196/16161](https://doi.org/10.2196/16161)] [Medline: [32406854](https://pubmed.ncbi.nlm.nih.gov/32406854/)]
42. Bakken S, Grullon-Figueroa L, Izquierdo R, Lee N, Morin P, Palmas W, IDEATel Consortium. Development, validation, and use of English and Spanish versions of the telemedicine satisfaction and usefulness questionnaire. *J Am Med Inform Assoc* 2006;13(6):660-667 [FREE Full text] [doi: [10.1197/jamia.M2146](https://doi.org/10.1197/jamia.M2146)] [Medline: [16929036](https://pubmed.ncbi.nlm.nih.gov/16929036/)]
43. Løchting I, Grotle M, Storheim K, Werner EL, Garratt AM. Individualized quality of life in patients with low back pain: reliability and validity of the Patient Generated Index. *J Rehabil Med* 2014 Sep;46(8):781-787 [FREE Full text] [doi: [10.2340/16501977-1826](https://doi.org/10.2340/16501977-1826)] [Medline: [25117176](https://pubmed.ncbi.nlm.nih.gov/25117176/)]
44. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 2002 Jan;82(1):8-24. [doi: [10.1093/ptj/82.1.8](https://doi.org/10.1093/ptj/82.1.8)] [Medline: [11784274](https://pubmed.ncbi.nlm.nih.gov/11784274/)]
45. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Harnessing different motivational frames via mobile phones to promote daily physical activity and reduce sedentary behavior in aging adults. *PLoS One* 2013;8(4):e62613 [FREE Full text] [doi: [10.1371/journal.pone.0062613](https://doi.org/10.1371/journal.pone.0062613)] [Medline: [23638127](https://pubmed.ncbi.nlm.nih.gov/23638127/)]

46. Mummah SA, Mathur M, King AC, Gardner CD, Sutton S. Mobile technology for vegetable consumption: a randomized controlled pilot study in overweight adults. *JMIR Mhealth Uhealth* 2016 May 18;4(2):e51 [FREE Full text] [doi: [10.2196/mhealth.5146](https://doi.org/10.2196/mhealth.5146)] [Medline: [27193036](https://pubmed.ncbi.nlm.nih.gov/27193036/)]
47. Melin J, Bonn SE, Pendrill L, Trolle Lagerros Y. A questionnaire for assessing user satisfaction with mobile health apps: development using Rasch measurement theory. *JMIR Mhealth Uhealth* 2020 May 26;8(5):e15909 [FREE Full text] [doi: [10.2196/15909](https://doi.org/10.2196/15909)] [Medline: [32452817](https://pubmed.ncbi.nlm.nih.gov/32452817/)]
48. Riis A, Rathleff MS, Hartvigsen J, Thomsen JL, Afzali T, Jensen MB. Feasibility study on recruitment in general practice for a low back pain online information study (part of the ADVIN Back Trial). *BMC Res Notes* 2020 Jan 10;13(1):24 [FREE Full text] [doi: [10.1186/s13104-020-4894-8](https://doi.org/10.1186/s13104-020-4894-8)] [Medline: [31924258](https://pubmed.ncbi.nlm.nih.gov/31924258/)]

Abbreviations

GRS: Global Rating Scale
ICC: intraclass correlation coefficient
LBP: low back pain
NPS: net promotor score
OPSI: Online Patient Satisfaction Index
REDCap: Research Electronic Data Capture
RMDQ: Roland-Morris Disability Questionnaire
SDCconsistency: smallest detectable change

Edited by R Kukafka, G Eysenbach; submitted 17.06.20; peer-reviewed by R Froud, F Bishop; comments to author 15.07.20; revised version received 07.09.20; accepted 29.09.21; published 15.11.21.

Please cite as:

Afzali T, Lauridsen HH, Thomsen JL, Hartvigsen J, Jensen MB, Riis A

The Online Patient Satisfaction Index for Patients With Low Back Pain: Development, Reliability, and Validation Study

JMIR Form Res 2021;5(11):e21462

URL: <https://formative.jmir.org/2021/11/e21462>

doi: [10.2196/21462](https://doi.org/10.2196/21462)

PMID: [34779785](https://pubmed.ncbi.nlm.nih.gov/34779785/)

©Tamana Afzali, Henrik Hein Lauridsen, Janus Laust Thomsen, Jan Hartvigsen, Martin Bach Jensen, Allan Riis. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 15.11.2021. 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

Game Plan, a Web-Based Intervention to Improve Adherence and Persistence to HIV Pre-exposure Prophylaxis and Reduce Heavy Drinking in Gay, Bisexual, and Other Men Who Have Sex With Men: Usability and User Experience Testing

Tyler B Wray¹, PhD; Philip A Chan², MD; Christopher W Kahler¹, PhD

¹Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, United States

²Department of Medicine, Warren Alpert Medical School of Brown University, Providence, RI, United States

Corresponding Author:

Tyler B Wray, PhD
Department of Behavioral and Social Sciences
Brown University School of Public Health
121 S Main Street
Box G-S121-5
Providence, RI, 02903
United States
Phone: 1 4018636659
Email: tyler_wray@brown.edu

Abstract

Background: Encouraging consistent use of pre-exposure prophylaxis (PrEP) is essential for reducing HIV incidence, particularly among gay, bisexual, and other men who have sex with men (GBM), and especially those who engage in heavy drinking. Although practice guidelines recommend providing adherence counseling to PrEP patients, clinics and providers may not have the resources or expertise to provide it. Internet-facilitated interventions have been shown to improve HIV prevention outcomes, including medication and care adherence. Game Plan is a website we created to help users make a tangible plan for reducing their HIV risk. We designed additional components of Game Plan to address key individual level barriers to PrEP use.

Objective: The aim of this mixed methods research is to test the usability and user experience of these components with intended users: GBM who drink heavily and are on PrEP.

Methods: In study 1 (usability), we completed a detailed individual interview in which participants (n=10) walked through a prototype of the website, *thinking aloud* as they did, and completed a follow-up interview and web-based survey afterward. Study 2 (user experience) involved providing participants (n=40) with a link to the prototype website to explore on their own and asking them to complete the same follow-up survey afterward. Qualitative data were analyzed using thematic analysis, and descriptive statistics were used to analyze quantitative data.

Results: Users in both studies gave the website excellent ratings for usability, overall satisfaction, and quality, and most often described the site as *informative*, *helpful*, and *supportive*. Users also rated the site's content and feel as respectful of them and their autonomy, empathetic, and they stated that it conveyed confidence in their ability to change. The study 1 interviews highlighted the importance of the website's esthetics to the participants' engagement with it and its credibility in prompting genuine reflection.

Conclusions: GBM who reported heavy drinking and used PrEP generally found a website focused on helping them to create a plan to use PrEP consistently to be helpful. Adopting user-centered design methods and attending to the esthetics of mobile health interventions are important steps toward encouraging engagement and reducing at-risk behaviors.

(*JMIR Form Res* 2021;5(11):e31843) doi:[10.2196/31843](https://doi.org/10.2196/31843)

KEYWORDS

pre-exposure prophylaxis; HIV; HIV prevention; mHealth; mobile health; eHealth; mobile phone

Introduction

Background

Medications approved for use as HIV pre-exposure prophylaxis (PrEP) have considerable promise of reducing overall HIV incidence [1,2], particularly among groups who are at a higher risk of HIV such as gay, bisexual, and other men who have sex with men (GBM) [3]. Despite an increase in PrEP use in recent years, uptake of PrEP for HIV prevention is still lower than levels needed to achieve a decline in new HIV diagnoses [4]. All PrEP medications currently approved in the United States, which include Truvada, Descovy, and the generic emtricitabine-tenofovir disoproxil fumarate (FTC-TDF), are recommended for once-daily oral dosing, in part because PrEP's efficacy in preventing HIV infection depends on consistent adherence [5]. Recent studies suggest that at least 4 doses per week provide approximately 96% efficacy [5,6]. Although many GBM using PrEP are able to achieve these levels of adherence [7-9], younger GBM may have high levels of suboptimal adherence, and they frequently discontinue using PrEP [9-12]. Interventions are needed to support GBM in successfully using PrEP while they are at risk.

Alcohol use is also an important risk factor for HIV acquisition [13], due in large part to alcohol's tendency to interfere with effective condom use during sex in GBM [14,15]. Therefore, GBM who drink heavily are key candidates for PrEP; yet, PrEP uptake among GBM who drink heavily may be lower than in more moderate drinkers [16]. Although there is little evidence that PrEP adherence is disrupted specifically when drinking [17,18], the personal characteristics of heavy drinkers or the lifestyle disruptions they experience may be associated with suboptimal adherence, and PrEP-alcohol interactive toxicity beliefs are common [16]. As such, interventions that address alcohol's potential role in PrEP uptake, adherence, and persistence would be valuable.

Current practice guidelines suggest that GBM follow up with their medical provider every 3 months while they are on PrEP for HIV and sexually transmitted infection (STI) testing and kidney function monitoring, among other things [19], but in practice, patients using PrEP often follow up less frequently [20,21]. Similarly, practice guidelines recommend providing adherence counseling to all patients using PrEP [19], but it is not clear how many patients using PrEP actually receive some form of counseling. It is likely that patients attending dedicated PrEP clinics are more likely to receive some form of adherence counseling, given that these clinics often have support staff available to provide it (eg, PrEP navigators) [22,23]. Although it is likely that a large proportion of PrEP patients receive their PrEP care through such specialty clinics, efforts to encourage more primary care providers to prescribe PrEP have been central to the effort to expand PrEP access [24]. The ability to provide support services such as adherence counseling is among the top barriers that primary care providers have identified to prescribing PrEP [25]. In patients using PrEP who do receive counseling, regardless of the setting, it is not clear to what extent the counseling uses techniques that are likely to improve

adherence or whether it adheres to the recommendations of current practice guidelines.

Meta-analyses have shown that digital intervention programs such as websites and smartphone apps can encourage HIV prevention behaviors in users [26,27]. Alongside traditional services, these programs could be particularly useful in helping improve PrEP outcomes, given that they are relatively inexpensive and easy to disseminate [28]. Another critical advantage of digital interventions is that they can also standardize the delivery of evidence-based behavior change techniques [29]. That is, the inherently programmatic delivery of digital interventions can expose users to a consistent set of content that uses specific techniques that have been shown to help change user behavior in past studies. In this way, digital interventions can ensure that all users are shown evidence-based content and can avoid problems with fidelity and program drift that normally affect interventions delivered by individuals. Although a variety of behavioral interventions have been developed to reinforce basic PrEP counseling and increase the success rate of GBM with PrEP, few have been tested in fully powered efficacy trials to date, and nearly all are intended for face-to-face delivery [23,30-38].

Game Plan is a website we originally developed to help GBM reduce their risk of HIV and alcohol use. It was designed to align with the *spirit* of motivational interviewing (MI) [39], with content informed by the information-motivation-behavior skills model [40] and using the basic flow of many brief motivational interventions [41]. An initial pilot trial showed promising results in reducing alcohol use and the number of anal sex partners in GBM [42]. We redesigned Game Plan and expanded its content to incorporate features to help promote PrEP uptake, adherence, and persistence, as well as to help PrEP users reduce their alcohol use and risk of STIs. Thorough descriptions of its content have been reported elsewhere (Wray, TB, unpublished data, April 2021) [43]. In brief, Game Plan provides specific, digestible feedback about users' risk of HIV given their behavior over the last year and shows them to what extent PrEP might reduce that risk with consistent adherence. It also presents information about users' risk of other STIs, the norms of users' sexual behaviors compared with those of other GBM in their age group, challenges several key *myths* about PrEP, and guides users through a decisional balance exercise to aid them in considering the pros and cons about their recent decisions about sex. Game Plan also guides users through various options for improving the consistency of their PrEP use and reducing their HIV and STI risk and suggests several practical steps that could help users accomplish these goals, similar to the change planning process in MI [39]. As was the case with the original Game Plan, this redesigned version provides personal feedback about users' recent alcohol use level and its potential health consequences, including specific information about how drinking heavily could interfere with PrEP adherence or persistence, among other things.

Objective

In this research, we employed user-centered design methods to test the usability [44] and overall user experience (UX) [45] of the redesigned Game Plan from the perspective of GBM who

were currently on PrEP. Thus, this research involved conducting 2 substudies: a usability study and a UX study. In the usability study (n=10), we conducted detailed individual interviews in which participants walked through a prototype version of the redesigned site to record their thoughts and reactions to the content and note any usability issues that arose. After the interview, these participants also completed a follow-up questionnaire to rate their perceptions of the site. In the UX study (n=40), participants were provided a link to the Game Plan site through email, and after finishing all portions of it, they were asked to complete the same follow-up survey that the usability participants completed. These procedures allowed us to examine how well the redesigned website met the goals of our design process, including whether GBM on PrEP perceived that the site (1) was easy to use, (2) provided content that prompted them to reflect on their HIV and STI risk, PrEP use, and drinking, (3) aligned with the *spirit* of MI, and (4) elicited high intentions to use the site under *real world* circumstances. Although Game Plan also incorporates an SMS text messaging feature that helps support users on their progress toward the goals they set on the site, this study focused primarily on the website interaction portion.

Methods

Participants

A total of 50 participants were recruited from gay-oriented dating apps (eg, Grindr and Scruff) and social networking websites (eg, Facebook and Instagram) in January-March 2021 to provide feedback on a staging version of Game Plan. Of the 50 participants, 10 (20%) completed the usability portion of the study, which involved completing a baseline survey, a 60- to 90-minute usability interview over videoconference, and a follow-up survey afterward. Of the 50 participants, 40 (80%) completed the UX portion, which involved completing the same surveys as the usability participants, but they completed the follow-up survey after perusing a staging version of the Game Plan site on their own. For both the usability and UX portions, eligible participants (1) were aged ≥ 18 years; (2) were cisgender men; (3) could speak and read fluently in English; (4) reported anal sex with a man in the past year; (5) self-reported as currently having a valid prescription for either Truvada, Descovy, or generic FTC-TDF as HIV PrEP; (6) were prescribed to take it daily, and (7) were classified as *hazardous drinkers* according to criteria from the National Institute on Alcohol Abuse and Alcoholism, that is, they reported having consumed either ≥ 14 drinks per week or ≥ 5 drinks on a single occasion at least once in the last month [46]. Demographic characteristics for all study participants are reported in Table 1.

Table 1. Demographic characteristics of the usability study participants (N=50).

Characteristics	Usability test (n=10)	UX ^a survey (n=40)
Age (range 21-68 years), mean (SD)	27.9 (3.1)	33.6 (10.9)
Race, n (%)		
White	7 (70)	34 (85)
Black or African American	1 (10)	2 (5)
American Indian or Alaska native	0 (0)	0 (0)
Asian	2 (20)	1 (3)
Pacific Islander or Native Hawaiian	0 (0)	0 (0)
Multiracial	0 (0)	2 (5)
Chose not to respond	0 (0)	1.0 (3)
Ethnicity (Hispanic or Latino), n (%)	1 (10)	9 (23)
Single relationship status, n (%)	7 (70)	31 (78)
College degree, n (%)	9 (90)	35 (88)
Low income ^a , n (%)	2 (20)	8 (20)
Unemployed, n (%)	1 (10)	1 (3)
Gay or bisexual identity, n (%)	10 (100)	39 (98)
US region of residence, n (%)		
Northeast	6 (60)	25 (63)
South	2 (20)	3 (8)
Midwest	1 (10)	6 (15)
West	1 (10)	6 (15)
AUDIT ^b total score, mean (SD)	12 (6.0)	10.8 (5.7)
DAST ^c total score, mean (SD)	1.7 (2.5)	1.7 (1.6)
PrEP^d medication type, n (%)		
Truvada	7 (70)	28 (70)
Descovy	3 (30)	10 (25)
FTC-TDF ^e	0 (0)	2 (5)
Number of years on PrEP, mean (SD)	4.0 (2.0)	4.7 (1.3)
Days taken PrEP, past 30 days, mean (SD)	26.1 (6.0)	28.0 (3.6)

^aUX: user experience.

^bAUDIT: Alcohol Use Disorders Identification Test.

^cDAST: Drug Abuse Screening Test.

^dPrEP: pre-exposure prophylaxis.

^eFTC-TDF: emtricitabine-tenofovir disoproxil fumarate.

Procedures

Participants interested in participating were provided a link to a screening survey through email. Those who met the eligibility criteria based on the screening survey then reviewed the web-based consent information in both the bullet-point format and in the approved consent document and, if interested, indicated their consent to participate in the full studies by clicking on a radio button. Participants who consented then provided basic contact information in a subsequent survey before completing an initial baseline survey that assessed basic

demographic information, relevant behavioral variables (eg, alcohol and drug use), and PrEP history. Afterward, participants in the usability portion of the project scheduled an appointment for an interview over videoconference. During this interview, participants were provided with a link to a staging version of the Game Plan site and asked to share their screen with the interviewers. They were then asked to use the site naturally and to *think aloud* as they did so. To ensure honest feedback, participants were also explicitly asked to be as critical and frank as they could be to provide us with realistic reactions. We collected video and audio recordings of the participants

themselves as well as their computer screens, to allow later review of comments, clicks and navigation, and facial expressions and reactions. After reviewing all the components, staff members then conducted a brief, semistructured interview inquiring about participants' perceptions of the app overall, including its language and tone, length and duration of its features, interactivity, ease of navigation, appropriateness, and most and least helpful components. These usability interviews lasted 60-90 minutes. Afterward, participants completed a brief web-based follow-up survey that assessed similar perceptions and reactions to the site. Participants in the UX study completed these same steps, with the exception of the usability interview portion. That is, they completed the same baseline survey but were asked to review the Game Plan site on their own, and they then completed the same follow-up survey as the usability participants. No revisions or changes to the site were made between the 2 studies. The usability study participants were compensated with US \$50 for completing all portions, and the UX study participants were compensated with US \$25. All procedures were approved by the Brown University Institutional Review Board.

Measures

PrEP History

Items in the baseline survey collected information about participants' PrEP use and history, including the medication they were currently taking (eg, Truvada, Descovy, or generic FTC-TDF), the date on which they were first prescribed PrEP, and self-reported daily adherence to PrEP over the last 30 days.

System Usability Scale

The System Usability Scale (SUS) [47] is a 10-item questionnaire that assesses a system's ease of use, intuitiveness, and desirability. Example items include the following: *I found the various functions in this system were well integrated* and *I think that I would like to use this system frequently*.

Participants rate each item on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*). The SUS has become an industry standard in the usability field, and it has excellent measurement properties ($\alpha=.90$) [48]. The SUS was collected from participants in both studies as part of the follow-up survey.

General Usability Items and Reactions

Participants also rated several items that asked about their reactions to the site generally as well as specific aspects of it. For example, items asked participants to rate how generally *engaging, interesting* and *boring* the site was on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*) and to rate the color scheme, font style, navigation, and organization of the site on a scale of 1 (*extremely disorganized* or *difficult* or *bad*) to 7 (*extremely organized* or *easy* or *good*). These items are shown in Table 2. Participants were also asked to pick up to 3 words from a list of 18 words that they believed best described Game Plan (eg, okay, helpful, terrible, all right, supportive, frustrating, unhelpful, and wonderful) and to pick and rank the components of the site that they found most and least helpful. Finally, they were also asked to rate how likely they believed that they would be to use Game Plan and its various components in their real lives (eg, email their change plan to themselves and sign up for SMS text messages to follow up on their goals) under various dissemination scenarios (eg, if a physician or a clinic provided them with a link or personally asked them to).

Table 2. Descriptive statistics for items rated by usability (n=10) and user experience study (n=40) participants (N=50).

Characteristics	Values, mean (SD)
Usability or esthetics^a	
Color scheme	6.0 (1.2)
Font style	6.6 (0.8)
Font size	6.6 (0.7)
Navigation use	6.0 (1.4)
Layout organization	6.3 (1.0)
Duration, tone, and engagement^b	
It was something I would want to use.	3.5 (1.1)
It was interesting.	4.1 (0.9)
It was engaging.	4.1 (0.8)
It was boring.	2.1 (0.9)
It took too long to use.	2.5 (1.1)
I had too much on my mind to really think about the content.	2.0 (1.0)
It was judgmental.	2.3 (1.3)
It helped me reflect on what I do in my own life.	3.9 (0.9)
The topics addressed were a good fit for me and other guys I know.	3.8 (0.9)
It was <i>sex negative</i> , or makes sex seem like a bad thing.	2.1 (1.2)
Motivational interviewing <i>spirit</i> items^b	
It provided feedback objectively, rather than trying to persuade me.	3.6 (1.1)
It told me what to do, rather than helping me understand what I want to do.	2.3 (1.0)
It was understanding of my point of view.	3.4 (1.2)
It showed me confidence in my ability to make changes.	3.7 (0.8)
It showed interest in my values and goals in life.	3.5 (1.0)
It tried to persuade me of the seriousness of the problem.	3.6 (1.0)
It encouraged me to contribute ideas about how to change my behavior.	3.4 (1.1)
It expressed disapproval of me.	2.2 (1.2)
It was empathetic, meaning that it showed an understanding of my perspective.	3.6 (1.0)
It showed respect for me.	4.1 (0.9)
It was supportive of my ability to decide for myself.	3.8 (0.8)
It was caring and friendly.	4.0 (0.8)
It was honest and trustworthy.	3.9 (0.9)

^aItems rated on a scale of 1 (*extremely bad or difficult or disorganized*) to 7 (*extremely good or easy or organized*).

^bItems were rated on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*).

Motivational Interviewing Spirit

A 13-item questionnaire was adapted from other common MI descriptions and rating systems [49-51] and was used to assess specific areas in which the app's content and approach were consistent or inconsistent with MI style. Each item (all of which are presented in Table 2) was rated on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*). In this study, we explored the summary statistics for individual items to determine specific areas for improvement that we might address in subsequent app versions.

The Mobile App Rating Scale: User Version

The Mobile App Rating Scale: User Version (uMARS) [52] is a 20-item rating scale that assesses the quality of mobile apps from the user's perspective. The uMARS asks participants to rate their perceptions of a specific app or site overall and across 4 subscales: engagement, functionality, esthetics, and information quality. In preliminary studies, the uMARS has shown excellent internal consistency ($\alpha=.90$) and good test-retest reliability [52]. The uMARS mean total score, subscale scores, and individual subjective quality ratings are reported here.

However, item 18 (*How many times do you think you would use this app in the next 12 months if it was relevant to you?*) was omitted from total scores because repeated use was not necessarily the primary use case for Game Plan. As total scores are simply participants' mean ratings on all assessed items, the impact of removing this item is likely to be minimal ($\alpha=.87$).

Motivation to Change Target Behaviors

Both the baseline and the follow-up surveys assessed participants' motivation to adhere to their PrEP and reduce their HIV and STI risk and alcohol use, using a single item reflecting participants' stages of change related to that behavior [53]. This question asks participants to rate which statement is most true for them, with 5 response options, including the following: *I don't see any need to [take my PrEP medication every day/reduce my drinking] and I've been [taking my PrEP medication every day/reducing my drinking] and will continue to do so.*

An additional question assessed to what extent participants would like to start changing each of these behaviors immediately, rating this on a scale of 1 (*not at all*) to 7 (*a bit*). Participants also reported whether they had made a plan to change each behavior (yes or no) and, if so, to rate how committed they were to following through with that plan on a scale of 1 (*not at all*) to 9 (*very much*).

Analysis

All interview audiotapes were transcribed, reviewed, and coded by 2 staff members using a thematic approach to analysis [54]. Given that the purpose of these interviews was to directly inform design decisions and revisions, the staff members specifically searched for key themes about the app's overall tone, interactivity, intuitiveness, and duration, as well as for specific feedback about the content on each screen. We also reviewed video clips of each session to note any significant facial expressions (eg, laughter, confusion, or frustration) and associated interaction behaviors (eg, clicks and data entry) that could further inform key themes. Once all transcripts were independently coded, staff members met to review them and develop consensus about consistent themes. To analyze quantitative data, we combined baseline and follow-up survey responses across both the usability and UX studies because the items assessed were the same and used descriptive statistics and pairwise 2-tailed *t* tests to compare specific items (eg, participants' intentions to use the app if provided a link by a clinic vs their physician asking them to use it directly). All analyzed data were used to guide design adjustments to the site before deployment in planned efficacy studies and to improve knowledge about web-based intervention design and dissemination (eg, components that GBM perceive to be most helpful and+ use intentions in various scenarios). Quantitative data were analyzed in Stata 16 (StataCorp), and all procedures were approved by the Brown University Institutional Review Board (protocol number 1905002453).

Results

Usability Interviews

Numerous Themes

The usability participants' specific suggestions on each page led to a number of important changes in Game Plan's interfaces. The postwalk-through semistructured interviews also generally suggested that the information and exercises provided through Game Plan were largely unique and not something many users had been exposed to before, despite all participants having significant experience with PrEP (mean 4.5, SD 1.5 years):

Overall, I really enjoyed it. I'd say that the...design was done really well and that sort of helped me through. There's a lot of information especially around alcohol and HIV and the other specifics of PrEP, especially things I didn't know. So I learned something new. [Participant 6]

These interviews yielded numerous other themes that could be important for PrEP and HIV interventions and, more generally, technology-facilitated interventions. Each of these themes is discussed below, with representative quotes from participants.

Theme 1: Aesthetics and Design Were Effective in Promoting Engagement

Many participants spontaneously noted that they thought the site was well-designed and esthetically pleasing, and several explicitly tied these characteristics to their engagement with the site:

It was definitely pretty. I liked the design, I liked the color scheme. I don't know that it was necessarily entertaining, but it was engaging the way the facts were presented along with it. I think that made it more attention-holding. [Participant 5]

Stylistically, it looks great. It drew me in. I think it would draw people in. I would probably have gone all the way through it even if I were just scrolling through on my own. [Participant 4]

Other participants suggested that although the site focused on health, it avoided an overly *clinical* look and incorporated an exciting and mysterious esthetic that would help pique the interests of users in this population:

The club/party design made it feel less medical and clinical. There's almost some sexual suggestiveness that's kind of intriguing about it. I think the sort of mysterious man in the background leads to a bit of intrigue and I think that's what would make me look through it. [Participant 6]

Theme 2: Credibility Is Critical for Sites That Focus on HIV and Other Health-Related Topics

Several participants commented on the importance of ensuring that a site that collected information about sexual behavior and provided information about HIV was credible. A participant noted that credibility was important to ensure that users would be willing to provide sensitive information about themselves,

although (as the site onboarding explicitly highlighted) the app collected no personal information and was anonymous:

Honestly, talking about sexual practices might seem like private information and I might be a little reluctant to share information, so I think it does help that there's like the Brown University branding on the first page, and that seems like a reputable institution so I feel comfortable moving forward. [Participant 10]

Other participants mentioned the importance of credibility specifically when viewing feedback about the potential health consequences of their recent behavior. They noted that providing trusted institution branding on the site and including footnotes citing credible sources alongside any key information or results might boost user confidence in the accuracy of that information:

I saw that there was the Brown University affiliation at the bottom, so for me, that makes me feel like everything in here is going to be correct, or at least extremely likely to be correct. It was also nice to also see these citations down here that make me think that it's more likely to be correct, although to be totally honest, I'll probably never actually look at them. [Participant 9]

Theme 3: Contrasting Users' HIV Risk While Taking PrEP Daily Versus Without PrEP Made a Strong Impression About the Value of PrEP Use

Several participants commented on how calculating their personal risk of HIV based on their recent behavior and then directly comparing it what their risk would be if they consistently used PrEP versus not using PrEP elicited a strong reaction that reinforced the importance of consistent PrEP adherence:

Even while I take PrEP, I still have some reservations and fears that I might contract HIV, but I didn't know it was that [emphasis theirs] effective. So that was a really reassuring number. [Participant 10]

The massive disparity between the [HIV] risk if you take [PrEP] and the risk if you don't, definitely was a bit of like... "Ok, I knew I should be doing it, but definitely if I miss it, I should be focusing more." [Participant 7]

These perspectives could support the notion that simply providing information about users' HIV risk level may be insufficient for increasing motivation to use PrEP. Instead, both correcting user misperceptions about their personal risk based on their recent behavior *and* showing them *to what extent* PrEP could reduce that risk with consistent adherence could increase PrEP use motivation more directly.

Theme 4: Feedback About Users' Personal Risk for STIs Given Recent Behavior Prompted Some PrEP Users to Consider Using Other Prevention Methods

Participants also noted that the specific feedback provided about their personal risk for other common STIs (gonorrhea and chlamydia) given their recent behavior also left an impression on them:

The Gonorrhea and Chlamydia numbers [also stood out to me]. It seems very high. Like I already knew it to be high, but knowing that I might contract Chlamydia from like 1 in 10 partners is pretty shocking. [Participant 10]

Other participants also explicitly noted considering using other forms of prevention (eg, condoms) after viewing the feedback about their personal risk for other STIs:

I think the numbers that I saw about Chlamydia and Gonorrhea might make me more inclined to use condoms. [Participant 3]

This theme suggests that providing similar feedback about users' personal risk for STIs other than HIV may leave a similarly strong impression on some PrEP users and prompt them to consider ways to reduce their STI risk.

Theme 5: Users Found Content Addressing Alcohol Use in an HIV and STI Intervention Useful Rather Than Intrusive

Several participants also noted that they found the site's content about alcohol use helpful, particularly because the role that drinking might play in increasing their sexual risk, potentially interfering with PrEP, was new to them:

The alcohol use section...I think was most helpful because it was tied to PrEP use and the risk of HIV. I had never made that kind of direct connection before. [Participant 10]

Others noted that although they agreed that alcohol-focused content was helpful in terms of considering alcohol's influence on their sexual decisions and PrEP use, they also believed it was helpful in reflecting on their alcohol use more generally. Specifically, a participant mentioned that the alcohol content helped them quantify their drinking pattern and reflect on its impact on their overall health and whether this pattern reflected their desired drinking level:

I think the alcohol section was the most helpful. I think it can be easy for me to, like, forget what the volumes that I drink are and how frequent they can become. I think understanding the volume of alcohol [was most helpful] and then also being reminded of the risks to like taking PrEP that can be associated with drinking. But I think I would probably still like to just in general try to reduce my alcohol use, even if there aren't benefits related to my sexual health. [Participant 9]

Overall, these themes largely support the conclusion that Game Plan's design and content were engaging, digestible, and prompted the sort of reflection in some target users that we had hoped for in creating Game Plan.

User Experience Survey

Participant ratings of individual items about the usability, esthetic characteristics, engagement, tone, and MI *spirit* are provided in [Table 2](#). The overall mean total score of the SUS was 83.1 (SD 14.2), which corresponds to an A grade (on a scale of A-F) [55] and is in the top 10% of scores, suggesting excellent usability and overall satisfaction [48]. The uMARS total mean

score of 4.0 (SD 0.5) further supports these findings, suggesting that users found the Game Plan site to be of above-average quality overall. Overall, participants gave the site 3.6 stars out of 5 (SD 0.8 stars) and stated that they would recommend the app to many people who might benefit from it (mean 3.3, SD 1.1).

Ratings of specific esthetic and design characteristics further support these overall findings, with participants rating Game Plan's color scheme, font style, font size, navigation, and layout and organization positively or very positively. Participant mean ratings on the functionality subscale of the uMARS, which primarily reflects usability-type concepts such as performance, ease of use, and navigation, was 4.2 (SD 0.7). Their ratings on the esthetics subscale of the uMARS, which reflects the general appearance of the app and its graphics, was also 4.2 (SD 0.6). Of the 50 participants, 38 (76%) chose positive words to describe Game Plan overall, with informative (76%), helpful (54%), and supportive (46%) being the 3 most commonly chosen overall. Of the 50 participants, 7 (14%) chose negative words, including *patronizing*, *wary*, and *judgmental*, showing some clear room for improvement. Together, these findings suggest that participants generally had positive reactions to the site, found it to be of high quality, and thought that it functioned smoothly and was visually appealing to them.

Participant ratings also generally suggested that the site's *tone* was generally at least somewhat consistent with our goals of conveying nonjudgmental, sex-positive content that helped them reflect on their own lives and was appropriate for them. Participant ratings also suggested that they perceived that Game Plan's content and tone were at least somewhat consistent with the core aspects of the MI *spirit*, including that it was respectful of them and their autonomy to decide for themselves, that it was caring and empathetic, and that it conveyed confidence in their ability to change. However, mean ratings of other elements were lower but acceptable, including that the site was understanding of the users' point of view and encouraged them to contribute ideas about how to change and suggesting some specific areas to improve upon.

When picking and ranking the sections of Game Plan that they found to be most helpful, participants most commonly included feedback about HIV risk (33/50, 66%), comparisons of social norms about sex (27/50, 54%), and feedback about PrEP's impact on HIV risk (22/40, 44%) to be among the 3 most helpful components, whereas they found the pros and cons (decisional balance) exercise (24/50, 48%), change planning for sexual risk reduction (22/50, 44%), and the feedback about STI risk (20/50, 40%) to be least helpful. However, participant mean ratings on the information quality subscale of the uMARS were generally high as well (mean 4.0, SD 0.6), suggesting that most participants found that the site had high quality, relevant information that was credible, not overwhelming, and presented very well visually.

Participants also generally rated Game Plan as somewhat engaging and interesting, with these items rated above the midpoint and a mean score on the uMARS engagement subscale of 3.5 (SD 0.6). Most users also agreed that they would want to use Game Plan if they were given access in the *real world*

(32/50, 64%), with 14% (7/50) suggesting that they would at least explore it. Furthermore, 46% (23/50) of the participants said that they would likely email their change plan from Game Plan to themselves afterward (13/50, 26%, possibly) and 28% (14/50) said that they would likely sign up for weekly SMS text messages to follow up on their progress toward their goals at the end of Game Plan (13/50, 26%, possibly). Interestingly, however, participant use intentions seemed to vary depending on how they were provided access to the site, with participants reporting that they believed they would be more likely to use Game Plan if their physician asked them to directly than if a clinic or physician's office simply provided them with a link to access it (mean 3.9, SD 1.2 vs mean 3.4, SD 1.3; $t_1=3.6$; $P<.001$). We also used pairwise t tests to explore whether the means of any ratings differed across minority participants versus non-Hispanic White participants or younger (aged <30 years) versus older participants, but no ratings were significantly different across these comparison groups.

Discussion

Principal Findings

In this study, we explored the usability and UX of a redesigned version of Game Plan among GBM who drink heavily and who were currently taking PrEP. Specific findings suggesting lack of clarity or usability problems with certain components led us to make many specific changes to the site design and interfaces before we began using Game Plan in 2 trials testing its efficacy in improving PrEP outcomes and HIV risk. For example, we revised the risk feedback to present it in a consistent format and direction (eg, HIV risk without PrEP and HIV risk with PrEP rather than to what extent PrEP reduced HIV risk) and changed how users entered certain data to increase ease of use (eg, the number of alcoholic drinks they consumed in the past 30 days), among other things. However, the results of these studies also produced important insights about internet-facilitated interventions more generally and their potential adoption among high-risk groups in the *real world*.

We identified 5 key themes from usability interviews that were frequently raised by participants. The first theme was that the site's esthetic characteristics contributed to user interest in it. Participants consistently noted their appreciation of the appearance of the site, and some explicitly noted that this likely increased their interest in the site, their attention to it, or their perseverance with it. Engagement, which is a form of attention, voluntary absorption, or curiosity often accompanied by motivation to use a specific product [56], is essential to the success of digital health interventions in changing behavior [57,58]. That is, users will not be exposed to the components of these interventions that actively contribute to behavior change if they are not first engaged. Several studies have shown that most users disengage from health behavior change technologies within the first few days in the *real world* [59,60], a significant problem for products that depend on repeated, even daily, use to effect change. Marketing research has shown that, in addition to ease of use and usefulness, user enjoyment of apps and websites is associated with their loyalty to them [61]. Furthermore, Cyr et al [62] found that users who rated the

appearance of apps as more pleasing also enjoyed these apps more, which in turn was associated with higher ratings of trust. As such, attending to the esthetic properties of websites and apps may both improve engagement and increase the credibility of its content.

The second theme further emphasized the importance of credibility and trustworthiness, especially for sites such as Game Plan that focus on topics such as HIV and other health issues. Participants noted that credibility of the content provided is essential if it is to influence user behavior. Research on alcohol interventions further supports the importance of credibility, particularly for techniques that provide users with feedback about the health consequences of behaviors or social approval of behaviors because the impact of these techniques on behavior change depends on users accepting and believing in the feedback provided [63]. Along with inattention, defensive reactions and doubt about the feedback could reduce the impact of this information on change [64-66]. Participants noted several aspects of the Game Plan design that seemed to increase its overall credibility, including prominently displaying the brand of a trusted institution and providing accessible references for key information wherever it appeared. These steps are consistent with past research showing that when evaluating the credibility of web-based information, users often check the sources of the information, who the author is, and the qualifications and credentials of the author when in doubt [67]. Designers of internet-facilitated health behavior change interventions should consider incorporating elements to enhance the credibility of feedback and information they provide users.

The third and fourth themes that emerged highlighted the value of providing digestible, specific feedback about user risk based on their recent behavior that prompted them to reflect on these decisions and the possibility of change. The third theme suggested that providing feedback to the user of the risk for HIV with and without PrEP side by side prompted participants to reflect on their recent adherence and commitment to taking PrEP every day. Past research has shown that many GBM are hesitant to start PrEP, in part because they may be underestimating their risk of HIV [68,69]. However, a brief intervention that used an HIV risk calculator to help correct recipient perceptions of their HIV risk did not seem to increase PrEP uptake in GBM [70]. Although there may be several reasons why this particular approach did not increase PrEP uptake, a possibility is that persuading GBM that they are underestimating their risk of HIV may highlight the need to take steps to reduce their risk but does not point to specific steps that would be successful. Our qualitative findings suggest that among PrEP users, illustrating the gap between their risk of HIV with and without PrEP use prompted several to reflect on the importance of consistent PrEP use. It is possible that this content could have a similar effect on GBM not yet using PrEP by illustrating to what extent they might reduce their risk of HIV if they started PrEP. Similarly, our findings show that some PrEP users noted that providing feedback about their risk of STIs other than HIV (eg, gonorrhea and chlamydia) also prompted them to consider taking other preventive steps (eg, using condoms and limiting sex partners) to reduce their risk. Encouraging PrEP users to reflect on their risk of other STIs

was a key goal in redesigning Game Plan, given the high rates of STIs in those on PrEP [71,72]. Despite the clear public health relevance of reducing STIs in those on PrEP, few interventions have addressed this issue to date.

The fifth theme suggested that many participants found the alcohol content useful, both in the context of reflecting on their HIV and STI risk and on its own. In particular, the participant responses indicated that many had not considered how their drinking might contribute to their risk of HIV or affect their success on PrEP, despite a history of regular interaction with sexual health services such as HIV and STI testing and PrEP care. These findings offer support for the notion that many GBM could be open to discussing their alcohol use when receiving sexual health services [73]. Some also specifically noted the value of reflecting on their alcohol use in the context of their general health, suggesting that GBM may appreciate discussing their drinking in general medical contexts as well. Importantly, there was no indication that addressing alcohol use in this context resulted in defensiveness that might have reduced engagement in Game Plan content around PrEP use.

Findings from the UX survey provided very strong support for the overall usability, design, and appeal of the Game Plan website. The overall SUS score of 83.1 is among the top 10% scores, suggesting exceptional usability and satisfaction. Past studies also suggest that products with scores this high are most likely to produce loyalty among users, and that users are most likely to recommend products with SUS scores >80 to their friends [74]. Mean ratings of participant intention to recommend the site to their friends further support this conclusion, with most responding that they would recommend it to many of their friends. Participant mean overall *star* rating was somewhat less enthusiastic, at 3.6 out of 5 (SD 0.8 stars). Similarly, participant ratings on markers of engagement were positive overall, with participants rating the site as at least *somewhat* interesting and engaging. Although these ratings suggest that there is room for improvement in these metrics, we believe that they are quite strong for a website that was explicitly designed to challenge user beliefs and perceptions and specifically intended to do so among users who were initially relatively indifferent toward change. Perhaps most importantly, however, 4 out of every 5 users reported that they believed they would be likely to at least explore Game Plan if provided access to it outside of the research context based on branding and marketing materials alone, suggesting that the website may be capable of achieving high engagement among GBM PrEP users in the absence of incentives or other motivations. For digital health interventions, demonstrating high use intention among intended users should be a prerequisite for moving beyond the design stage.

Finally, GBM PrEP user ratings of Game Plan's overall *tone* and alignment with the *spirit* of MI were also very encouraging and suggest that we largely met our design goals. Participant mean ratings of these items suggested that they agreed at least *somewhat* that Game Plan showed respect for them and was collaborative, empathetic, and nonjudgmental. Ensuring that Game Plan design and content aligned with these principles was essential, given that meta-analyses have identified expressing empathy and other aspects of the MI spirit (eg, collaboration and respect for autonomy) as some of the strongest

potential mechanisms of MI counseling effects on health behavior change [75].

Limitations

Although this study provides strong preliminary evidence of Game Plan's overall usability and high engagement, several limitations should be noted. First, although the sample sizes of both the usability and UX surveys were larger than those of many other similar studies, larger sample sizes would likely yield more confidence in these findings. Second, this sample consisted largely of White, non-Hispanic, well-educated, and higher-income GBM; therefore, the results reported may differ in a more diverse sample. Third, participants were considered eligible if they self-reported currently having a valid PrEP prescription; therefore, it is possible that some participants were not on PrEP at the time of their participation. Fourth, participants were also very highly adherent to it over the past 30 days. Although Game Plan was explicitly redesigned to incorporate content to help PrEP users be successful in PrEP care, PrEP users likely inherently represent a subset of GBM who are especially motivated to reduce their risk for HIV. As such, they may be more likely to provide positive ratings in response to

web-based tools that help them consider ways of reducing their HIV and STI risk, especially when compared with GBM who are less motivated to do so.

Overall, this study showed that GBM who use PrEP found the Game Plan website easy to use and engaging, and most of them reported that they would be likely to use it if a clinic or their physician provided them access to it in their real lives. The findings also suggested that the site's overall branding and esthetic characteristics contributed to participants' motivation to use it. The results also showed that specific Game Plan content seemed to be encouraging users to reflect on many of the issues we had intended: the value of consistent PrEP adherence for HIV risk reduction, users' continued risk of STIs, and alcohol's potential contribution to these issues. Now that Game Plan's basic usability and potential for broad uptake have been established, future research is needed to test its efficacy in encouraging key outcomes in PrEP care, such as consistent PrEP adherence during periods of risk and reducing other risk factors of HIV acquisition, including STIs and heavy drinking. We are initiating trials to explore these outcomes in the coming months.

Acknowledgments

This study was supported by research project grants R34AA027195, P01AA019072, P30AI042853, and loan repayment program grant L30AA023336 from the National Institute on Alcohol Abuse and Alcoholism.

Conflicts of Interest

None declared.

References

1. Grant RM, Anderson PL, McMahan V, Liu A, Amico KR, Mehrotra M, et al. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. *Lancet Infect Dis* 2014 Sep;14(9):820-829. [doi: [10.1016/s1473-3099\(14\)70847-3](https://doi.org/10.1016/s1473-3099(14)70847-3)]
2. Liu AY, Cohen SE, Vittinghoff E, Anderson PL, Doblecki-Lewis S, Bacon O, et al. Preexposure prophylaxis for HIV infection integrated with municipal- and community-based sexual health services. *JAMA Intern Med* 2016 Jan 01;176(1):75-84 [FREE Full text] [doi: [10.1001/jamainternmed.2015.4683](https://doi.org/10.1001/jamainternmed.2015.4683)] [Medline: [26571482](https://pubmed.ncbi.nlm.nih.gov/26571482/)]
3. HIV Surveillance Report, 2018 (Updated). Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2018-updated-vol-31.pdf> [accessed 2021-10-11]
4. LeVasseur MT, Goldstein ND, Tabb LP, Olivieri-Mui BL, Welles SL. The effect of PrEP on HIV incidence among men who have sex with men in the context of condom use, treatment as prevention, and seroadaptive practices. *J Acquir Immune Defic Syndr* 2018 Jan 01;77(1):31-40. [doi: [10.1097/QAI.0000000000001555](https://doi.org/10.1097/QAI.0000000000001555)] [Medline: [28961679](https://pubmed.ncbi.nlm.nih.gov/28961679/)]
5. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, iPrEx Study Team. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. *Sci Transl Med* 2012 Sep 12;4(151):151ra125 [FREE Full text] [doi: [10.1126/scitranslmed.3004006](https://doi.org/10.1126/scitranslmed.3004006)] [Medline: [22972843](https://pubmed.ncbi.nlm.nih.gov/22972843/)]
6. Buchbinder SP. Maximizing the benefits of HIV preexposure prophylaxis. *Top Antivir Med* 2018 Apr;25(4):138-142 [FREE Full text] [Medline: [29689539](https://pubmed.ncbi.nlm.nih.gov/29689539/)]
7. van Epps P, Maier M, Lund B, Howren M, Beck B, Beste L, et al. Medication adherence in a nationwide cohort of veterans initiating pre-exposure prophylaxis (PREP) to prevent HIV infection. *J Acquir Immune Defic Syndr* 2018 Mar 01;77(3):272-278. [doi: [10.1097/QAI.0000000000001598](https://doi.org/10.1097/QAI.0000000000001598)] [Medline: [29210835](https://pubmed.ncbi.nlm.nih.gov/29210835/)]
8. Marcus J, Hurley L, Hare C, Nguyen D, Phengrasamy T, Silverberg M, et al. Preexposure prophylaxis for HIV prevention in a large integrated health care system: adherence, renal safety, and discontinuation. *J Acquir Immune Defic Syndr* 2016 Dec 15;73(5):540-546 [FREE Full text] [doi: [10.1097/QAI.0000000000001129](https://doi.org/10.1097/QAI.0000000000001129)] [Medline: [27851714](https://pubmed.ncbi.nlm.nih.gov/27851714/)]
9. Hosek S, Rudy B, Landovitz R, Kapogiannis B, Siberry G, Rutledge B, Adolescent Trials Network (ATN) for HIVAIDS Interventions. An HIV preexposure prophylaxis demonstration project and safety study for young MSM. *J Acquir Immune Defic Syndr* 2017 Jan 01;74(1):21-29 [FREE Full text] [doi: [10.1097/QAI.0000000000001179](https://doi.org/10.1097/QAI.0000000000001179)] [Medline: [27632233](https://pubmed.ncbi.nlm.nih.gov/27632233/)]

10. Giler R, Magnuson D, Hawkins T, Bush S, Rawlings K, McCallister S. Changes in Truvada® for HIV pre-exposure prophylaxis utilization in the USA: 2012-2016 - Disparities. In: Proceedings of the 19th IAS Conference on HIV Science. 2017 Presented at: 19th IAS Conference on HIV Science; July 23-26, 2017; Paris, France URL: https://www.natap.org/2017/IAS/IAS_72.htm
11. Morgan E, Ryan DT, Newcomb ME, Mustanski B. High rate of discontinuation may diminish PrEP coverage among young men who have sex with men. *AIDS Behav* 2018 Nov 4;22(11):3645-3648 [FREE Full text] [doi: [10.1007/s10461-018-2125-2](https://doi.org/10.1007/s10461-018-2125-2)] [Medline: [29728950](https://pubmed.ncbi.nlm.nih.gov/29728950/)]
12. Hosek SG, Landovitz RJ, Kapogiannis B, Siberry GK, Rudy B, Rutledge B, et al. Safety and feasibility of antiretroviral preexposure prophylaxis for adolescent men who have sex with men aged 15 to 17 years in the United States. *JAMA Pediatr* 2017 Nov 01;171(11):1063-1071 [FREE Full text] [doi: [10.1001/jamapediatrics.2017.2007](https://doi.org/10.1001/jamapediatrics.2017.2007)] [Medline: [28873128](https://pubmed.ncbi.nlm.nih.gov/28873128/)]
13. Sander P, Cole S, Stall R, Jacobson L, Eron J, Napravnik S, et al. Joint effects of alcohol consumption and high-risk sexual behavior on HIV seroconversion among men who have sex with men. *AIDS* 2013 Mar 13;27(5):815-823 [FREE Full text] [doi: [10.1097/QAD.0b013e32835cff4b](https://doi.org/10.1097/QAD.0b013e32835cff4b)] [Medline: [23719351](https://pubmed.ncbi.nlm.nih.gov/23719351/)]
14. Wray TB, Celio MA, Pérez AE, DiGiuseppi GT, Carr DJ, Woods LA, et al. Causal effects of alcohol intoxication on sexual risk intentions and condom negotiation skills among high-risk men who have sex with men (MSM). *AIDS Behav* 2019 Jan 7;23(1):161-174 [FREE Full text] [doi: [10.1007/s10461-018-2243-x](https://doi.org/10.1007/s10461-018-2243-x)] [Medline: [30088199](https://pubmed.ncbi.nlm.nih.gov/30088199/)]
15. Shuper PA, Joharchi N, Monti PM, Loutfy M, Rehm J. Acute alcohol consumption directly increases hiv transmission risk: a randomized controlled experiment. *J Acquir Immune Defic Syndr* 2017 Dec 15;76(5):493-500 [FREE Full text] [doi: [10.1097/QAI.0000000000001549](https://doi.org/10.1097/QAI.0000000000001549)] [Medline: [28930769](https://pubmed.ncbi.nlm.nih.gov/28930769/)]
16. Kalichman SC, Eaton L. Alcohol-antiretroviral interactive toxicity beliefs as a potential barrier to HIV pre-exposure prophylaxis among men who have sex with men. *J Int AIDS Society* 2017 Jul 17;20(1):21534. [doi: [10.7448/ias.20.1.21534](https://doi.org/10.7448/ias.20.1.21534)]
17. Wray T, Chan P, Kahler C, Simpanen E, Liu T, Mayer K. Vulnerable periods: characterizing patterns of sexual risk and substance use during lapses in adherence to HIV pre-exposure prophylaxis among men who have sex with men. *J Acquir Immune Defic Syndr* 2019 Mar 01;80(3):276-283 [FREE Full text] [doi: [10.1097/QAI.0000000000001914](https://doi.org/10.1097/QAI.0000000000001914)] [Medline: [30531302](https://pubmed.ncbi.nlm.nih.gov/30531302/)]
18. Oldfield BJ, Edelman EJ. Addressing unhealthy alcohol use and the HIV pre-exposure prophylaxis care continuum in primary care: a scoping review. *AIDS Behav* 2021 Jun 20;25(6):1777-1789. [doi: [10.1007/s10461-020-03107-6](https://doi.org/10.1007/s10461-020-03107-6)] [Medline: [33219492](https://pubmed.ncbi.nlm.nih.gov/33219492/)]
19. Preexposure prophylaxis for the prevention of HIV infection in the United States - 2017 update : a clinical practice guideline. Centers for Disease Control and Prevention: US Public Health Service. 2018. URL: <https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2017.pdf> [accessed 2021-10-11]
20. Chan PA, Mena L, Patel R, Oldenburg CE, Beauchamps L, Perez-Brumer AG, et al. Retention in care outcomes for HIV pre-exposure prophylaxis implementation programmes among men who have sex with men in three US cities. *J Int AIDS Soc* 2016;19(1):20903 [FREE Full text] [doi: [10.7448/IAS.19.1.20903](https://doi.org/10.7448/IAS.19.1.20903)] [Medline: [27302837](https://pubmed.ncbi.nlm.nih.gov/27302837/)]
21. Chan PA, Patel RR, Mena L, Marshall BD, Rose J, Suttan Coats C, et al. Long-term retention in pre-exposure prophylaxis care among men who have sex with men and transgender women in the United States. *J Int AIDS Soc* 2019 Aug;22(8):e25385 [FREE Full text] [doi: [10.1002/jia2.25385](https://doi.org/10.1002/jia2.25385)] [Medline: [31423756](https://pubmed.ncbi.nlm.nih.gov/31423756/)]
22. Calabrese SK, Magnus M, Mayer KH, Krakower DS, Eldahan AI, Hawkins LA, et al. Putting PrEP into practice: lessons learned from early-adopting U.S. Providers' firsthand experiences providing HIV pre-exposure prophylaxis and associated care. *PLoS One* 2016 Jun 15;11(6):e0157324 [FREE Full text] [doi: [10.1371/journal.pone.0157324](https://doi.org/10.1371/journal.pone.0157324)] [Medline: [27304883](https://pubmed.ncbi.nlm.nih.gov/27304883/)]
23. Ming K, Shrestha I, Vazquez A, Wendelborn J, Jimenez V, Lisha N, et al. Improving the HIV PrEP continuum of care using an intervention for healthcare providers: a stepped-wedge study protocol. *BMJ Open* 2020 Jul 14;10(7):e040734 [FREE Full text] [doi: [10.1136/bmjopen-2020-040734](https://doi.org/10.1136/bmjopen-2020-040734)] [Medline: [32665393](https://pubmed.ncbi.nlm.nih.gov/32665393/)]
24. Calabrese SK, Krakower DS, Mayer KH. Integrating HIV Preexposure Prophylaxis (PrEP) into routine preventive health care to avoid exacerbating disparities. *Am J Public Health* 2017 Dec;107(12):1883-1889. [doi: [10.2105/ajph.2017.304061](https://doi.org/10.2105/ajph.2017.304061)]
25. Pleuhs B, Quinn KG, Walsh JL, Petroll AE, John SA. Health care provider barriers to HIV pre-exposure prophylaxis in the United States: a systematic review. *AIDS Patient Care STDS* 2020 Mar;34(3):111-123 [FREE Full text] [doi: [10.1089/apc.2019.0189](https://doi.org/10.1089/apc.2019.0189)] [Medline: [32109141](https://pubmed.ncbi.nlm.nih.gov/32109141/)]
26. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. *J Med Internet Res* 2014 May 26;16(5):e134 [FREE Full text] [doi: [10.2196/jmir.3393](https://doi.org/10.2196/jmir.3393)] [Medline: [24862459](https://pubmed.ncbi.nlm.nih.gov/24862459/)]
27. Noar S, Black H, Pierce L. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. *AIDS* 2009 Jan 02;23(1):107-115. [doi: [10.1097/QAD.0b013e32831c5500](https://doi.org/10.1097/QAD.0b013e32831c5500)] [Medline: [19050392](https://pubmed.ncbi.nlm.nih.gov/19050392/)]
28. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. *Am J Prev Med* 2013 Aug;45(2):228-236 [FREE Full text] [doi: [10.1016/j.amepre.2013.03.017](https://doi.org/10.1016/j.amepre.2013.03.017)] [Medline: [23867031](https://pubmed.ncbi.nlm.nih.gov/23867031/)]
29. Miller WR, Sorensen JL, Selzer JA, Brigham GS. Disseminating evidence-based practices in substance abuse treatment: a review with suggestions. *J Subst Abuse Treat* 2006 Jul;31(1):25-39. [doi: [10.1016/j.jsat.2006.03.005](https://doi.org/10.1016/j.jsat.2006.03.005)] [Medline: [16814008](https://pubmed.ncbi.nlm.nih.gov/16814008/)]

30. Golub S, Peña S, Hilley A, Pachankis J, Radix A. Brief behavioral intervention increases PrEP drug levels in a real-world setting. In: Proceedings of the Conference on Retroviruses and Opportunistic Infections. 2017 Presented at: Conference on Retroviruses and Opportunistic Infections; February 13-16, 2017; Seattle, Washington URL: <https://www.croiconference.org/abstract/brief-behavioral-intervention-increases-prep-drug-levels-real-world-setting/>
31. Reback CJ, Clark KA, Runger D, Fehrenbacher AE. A promising PrEP navigation intervention for transgender women and men who have sex with men experiencing multiple syndemic health disparities. *J Community Health* 2019 Dec 17;44(6):1193-1203 [FREE Full text] [doi: [10.1007/s10900-019-00705-x](https://doi.org/10.1007/s10900-019-00705-x)] [Medline: [31317438](https://pubmed.ncbi.nlm.nih.gov/31317438/)]
32. Amico KR, Miller J, Balthazar C, Serrano PA, Brothers J, Zollweg S, et al. Integrated Next Step Counseling (iNSC) for sexual health and prep use among young men who have sex with men: implementation and observations from ATN110/113. *AIDS Behav* 2019 Jul 9;23(7):1812-1823 [FREE Full text] [doi: [10.1007/s10461-018-2291-2](https://doi.org/10.1007/s10461-018-2291-2)] [Medline: [30315429](https://pubmed.ncbi.nlm.nih.gov/30315429/)]
33. Raifman J, Nunn A, Oldenburg CE, Montgomery MC, Almonte A, Agwu AL, et al. An evaluation of a clinical pre-exposure prophylaxis education intervention among men who have sex with men. *Health Serv Res* 2018 Aug 25;53(4):2249-2267 [FREE Full text] [doi: [10.1111/1475-6773.12746](https://doi.org/10.1111/1475-6773.12746)] [Medline: [28744983](https://pubmed.ncbi.nlm.nih.gov/28744983/)]
34. Fuchs JD, Stojanovski K, Vittinghoff E, McMahan VM, Hosek SG, Amico KR, et al. A mobile health strategy to support adherence to antiretroviral preexposure prophylaxis. *AIDS Patient Care STDS* 2018 Mar;32(3):104-111 [FREE Full text] [doi: [10.1089/apc.2017.0255](https://doi.org/10.1089/apc.2017.0255)] [Medline: [29565183](https://pubmed.ncbi.nlm.nih.gov/29565183/)]
35. Young LE, Schumm P, Alon L, Bouris A, Ferreira M, Hill B, et al. PrEP Chicago: a randomized controlled peer change agent intervention to promote the adoption of pre-exposure prophylaxis for HIV prevention among young Black men who have sex with men. *Clin Trials* 2018 Feb 11;15(1):44-52 [FREE Full text] [doi: [10.1177/1740774517730012](https://doi.org/10.1177/1740774517730012)] [Medline: [28862483](https://pubmed.ncbi.nlm.nih.gov/28862483/)]
36. Patel VV, Ginsburg Z, Golub SA, Horvath KJ, Rios N, Mayer KH, et al. Empowering with PrEP (E-PrEP), a peer-led social media-based intervention to facilitate HIV preexposure prophylaxis adoption among young black and latinx gay and bisexual men: protocol for a cluster randomized controlled trial. *JMIR Res Protoc* 2018 Aug 28;7(8):e11375 [FREE Full text] [doi: [10.2196/11375](https://doi.org/10.2196/11375)] [Medline: [30154071](https://pubmed.ncbi.nlm.nih.gov/30154071/)]
37. Moitra E, van den Berg JJ, Sowemimo-Coker G, Chau S, Nunn A, Chan PA. Open pilot trial of a brief motivational interviewing-based HIV pre-exposure prophylaxis intervention for men who have sex with men: preliminary effects, and evidence of feasibility and acceptability. *AIDS Care* 2020 Mar 25;32(3):406-410 [FREE Full text] [doi: [10.1080/09540121.2019.1622644](https://doi.org/10.1080/09540121.2019.1622644)] [Medline: [31130000](https://pubmed.ncbi.nlm.nih.gov/31130000/)]
38. Kelly JA, Amirkhanian YA, Walsh JL, Brown KD, Quinn KG, Petroll AE, et al. Social network intervention to increase pre-exposure prophylaxis (PrEP) awareness, interest, and use among African American men who have sex with men. *AIDS Care* 2020 May 13;32(sup2):40-46 [FREE Full text] [doi: [10.1080/09540121.2020.1739207](https://doi.org/10.1080/09540121.2020.1739207)] [Medline: [32167374](https://pubmed.ncbi.nlm.nih.gov/32167374/)]
39. Miller W, Rollnick S. *Motivational Interviewing: Helping People Change*. New York: Guilford Press; 2012.
40. Fisher JD, Fisher WA, Bryan AD, Misovich SJ. Information-motivation-behavioral skills model-based HIV risk behavior change intervention for inner-city high school youth. *Health Psychol* 2002;21(2):177-186. [doi: [10.1037/0278-6133.21.2.177](https://doi.org/10.1037/0278-6133.21.2.177)]
41. Field C, Hungerford D, Dunn C. Brief motivational interventions: an introduction. *J Trauma* 2005 Sep;59(3 Suppl):21-26. [doi: [10.1097/01.ta.0000179899.37332.8a](https://doi.org/10.1097/01.ta.0000179899.37332.8a)] [Medline: [16355056](https://pubmed.ncbi.nlm.nih.gov/16355056/)]
42. Wray TB, Kahler CW, Simpanen EM, Operario D. A preliminary randomized controlled trial of game plan, a web application to help men who have sex with men reduce their HIV risk and alcohol use. *AIDS Behav* 2019 Jun 22;23(6):1668-1679 [FREE Full text] [doi: [10.1007/s10461-019-02396-w](https://doi.org/10.1007/s10461-019-02396-w)] [Medline: [30671682](https://pubmed.ncbi.nlm.nih.gov/30671682/)]
43. Wray T, Kahler CW, Simpanen EM, Operario D. Game plan: development of a web app designed to help men who have sex with men reduce their HIV risk and alcohol use. *JMIR Form Res* 2018 Aug 23;2(2):e10125 [FREE Full text] [doi: [10.2196/10125](https://doi.org/10.2196/10125)] [Medline: [30684415](https://pubmed.ncbi.nlm.nih.gov/30684415/)]
44. Goodman E, Kuniavsky M, Moed A. *Observing the User Experience: A Practitioner's Guide to User Research*. Amsterdam: Elsevier; 2012.
45. Albert W, Tullis T, Tedesco D. *Beyond the Usability Lab: Conducting Large-scale Online User Experience Studies*. Burlington, Massachusetts, United States: Morgan Kaufmann; 2010.
46. *Helping patients who drink too much: a clinician's guide*. U.S. Department of Health and Human Services, National Institutes of Health, and National Institute on Alcohol Abuse and Alcoholism. 2005. URL: <http://www.webcitation.org/6vAQJtnqO> [accessed 2017-11-02]
47. Brooke J. SUS-A quick and dirty usability scale. In: *Usability Evaluation in Industry*. Boca Raton, Florida, United States: CRC Press; 1996:4-7.
48. Lewis J, Sauro J. The factor structure of the System Usability Scale. In: *Proceedings of the International Conference on Human Centered Design*. 2009 Presented at: International Conference on Human Centered Design; July 19-24, 2009; San Diego, CA, USA. [doi: [10.1007/978-3-642-02806-9_12](https://doi.org/10.1007/978-3-642-02806-9_12)]
49. Substance Abuse and Mental Health Services Administration (US). Chapter 3—Motivational interviewing as a counseling style. In: *Enhancing Motivation for Change in Substance Use Disorder Treatment*. Rockville (MD): Substance Abuse and Mental Health Services Administration (US); 2019.

50. Moyers T, Martin T, Manuel J, Miller W, Ernst D. Revised global scales: motivational interviewing treatment integrity 3.1.1 (MITI 3.1.1). Unpublished Manuscript, University of New Mexico, Albuquerque, NM. 2010. URL: https://casaa.unm.edu/download/miti3_1.pdf [accessed 2021-10-11]
51. Miller W, Moyers T, Ernst D, Amrhein P. Manual for the motivational interviewing skill code (MISC). Unpublished manuscript. Albuquerque: Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico. 2003. URL: <https://digitalcommons.montclair.edu/psychology-facpubs/27/> [accessed 2021-10-11]
52. 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/)]
53. Cook CL, Perri M. Single-item vs multiple-item measures of stage of change in compliance with prescribed medications. *Psychol Rep* 2004 Feb 01;94(1):115-124. [doi: [10.2466/pr0.94.1.115-124](https://doi.org/10.2466/pr0.94.1.115-124)] [Medline: [15077755](https://pubmed.ncbi.nlm.nih.gov/15077755/)]
54. Braun V, Clarke V. *Thematic Analysis*. New York: Springer; 2014.
55. Sauro J. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Denver, CO: Measuring Usability LLC; 2011:1-162.
56. Chapman P. *Models of Engagement : Intrinsically Motivated Interaction With Multimedia Learning Software*. Waterloo, Ontario: University of Waterloo; 1997.
57. 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/)]
58. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med* 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: [10.1007/s13142-016-0453-1](https://doi.org/10.1007/s13142-016-0453-1)] [Medline: [27966189](https://pubmed.ncbi.nlm.nih.gov/27966189/)]
59. Fox G, Garland S, Keibel A, Saxon LA. Why people stick with or abandon wearable devices. *NEJM Catalyst* 2017;3(5):0396 [FREE Full text] [doi: [10.1056/CAT.17.0396](https://doi.org/10.1056/CAT.17.0396)]
60. 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/)]
61. Kumar DS, Purani K, Viswanathan SA. Influences of ‘appscape’ on mobile app adoption and m-loyalty. *J Retail Consum Serv* 2018 Nov;45:132-141. [doi: [10.1016/j.jretconser.2018.08.012](https://doi.org/10.1016/j.jretconser.2018.08.012)]
62. Cyr D, Head M, Ivanov A. Design aesthetics leading to m-loyalty in mobile commerce. *Inform Manag* 2006 Dec;43(8):950-963. [doi: [10.1016/j.im.2006.08.009](https://doi.org/10.1016/j.im.2006.08.009)]
63. Boyle SC, Earle AM, LaBrie JW, Smith DJ. PNF 2.0? Initial evidence that gamification can increase the efficacy of brief, web-based personalized normative feedback alcohol interventions. *Addict Behav* 2017 Apr;67:8-17 [FREE Full text] [doi: [10.1016/j.addbeh.2016.11.024](https://doi.org/10.1016/j.addbeh.2016.11.024)] [Medline: [27978426](https://pubmed.ncbi.nlm.nih.gov/27978426/)]
64. Hummer JF, Davison GC. Examining the role of source credibility and reference group proximity on personalized normative feedback interventions for college student alcohol use: a randomized laboratory experiment. *Subst Use Misuse* 2016 Nov 09;51(13):1701-1715. [doi: [10.1080/10826084.2016.1197258](https://doi.org/10.1080/10826084.2016.1197258)] [Medline: [27487077](https://pubmed.ncbi.nlm.nih.gov/27487077/)]
65. LaBrie JW, Hummer JF, Neighbors C, Larimer ME. Whose opinion matters? The relationship between injunctive norms and alcohol consequences in college students. *Addict Behav* 2010 Apr;35(4):343-349 [FREE Full text] [doi: [10.1016/j.addbeh.2009.12.003](https://doi.org/10.1016/j.addbeh.2009.12.003)] [Medline: [20045262](https://pubmed.ncbi.nlm.nih.gov/20045262/)]
66. Lewis MA, Neighbors C. An examination of college student activities and attentiveness during a web-delivered personalized normative feedback intervention. *Psychol Addict Behav* 2015 Mar;29(1):162-167 [FREE Full text] [doi: [10.1037/adb0000003](https://doi.org/10.1037/adb0000003)] [Medline: [25134036](https://pubmed.ncbi.nlm.nih.gov/25134036/)]
67. Hargittai E, Fullerton L, Menchen-Trevino E, Thomas K. Trust online: young adults' evaluation of web content. *Int J Commun* 2010;4:27 [FREE Full text]
68. Gallagher T, Link L, Ramos M, Bottger E, Aberg J, Daskalakis D. Self-perception of HIV risk and candidacy for pre-exposure prophylaxis among men who have sex with men testing for HIV at commercial sex venues in New York city. *LGBT Health* 2014 Sep;1(3):218-224. [doi: [10.1089/lgbt.2013.0046](https://doi.org/10.1089/lgbt.2013.0046)] [Medline: [26789715](https://pubmed.ncbi.nlm.nih.gov/26789715/)]
69. Oostrom L, Rosentel K, Motley D, Hill BJ. Discordance in objective and self-perceived HIV risk: a potential barrier to pre-exposure prophylaxis in young gay and bisexual men. *J Assoc Nurses AIDS Care* 2020;31(1):103-109. [doi: [10.1097/jnc.000000000000137](https://doi.org/10.1097/jnc.000000000000137)]
70. Blumenthal J, Jain S, Mulvihill E, Sun S, Hanashiro M, Ellorin E, et al. Perceived versus calculated HIV risk: implications for pre-exposure prophylaxis uptake in a randomized trial of men who have sex with men. *J Acquir Immune Defic Syndr* 2019 Feb 01;80(2):23-29 [FREE Full text] [doi: [10.1097/QAI.0000000000001888](https://doi.org/10.1097/QAI.0000000000001888)] [Medline: [30422909](https://pubmed.ncbi.nlm.nih.gov/30422909/)]
71. Traeger MW, Cornelisse VJ, Asselin J, Price B, Roth NJ, Willcox J, PrEPX Study Team. Association of HIV preexposure prophylaxis with incidence of sexually transmitted infections among individuals at high risk of HIV infection. *J Am Med Assoc* 2019 Apr 09;321(14):1380-1390 [FREE Full text] [doi: [10.1001/jama.2019.2947](https://doi.org/10.1001/jama.2019.2947)] [Medline: [30964528](https://pubmed.ncbi.nlm.nih.gov/30964528/)]

72. Werner RN, Gaskins M, Nast A, Dressler C. Incidence of sexually transmitted infections in men who have sex with men and who are at substantial risk of HIV infection – A meta-analysis of data from trials and observational studies of HIV pre-exposure prophylaxis. PLoS ONE 2018 Dec 3;13(12):e0208107. [doi: [10.1371/journal.pone.0208107](https://doi.org/10.1371/journal.pone.0208107)]
73. Wray TB, Grin B, Dorfman L, Glynn TR, Kahler CW, Marshall BD, et al. Systematic review of interventions to reduce problematic alcohol use in men who have sex with men. Drug Alcohol Rev 2016 Mar 13;35(2):148-157 [FREE Full text] [doi: [10.1111/dar.12271](https://doi.org/10.1111/dar.12271)] [Medline: [25866929](https://pubmed.ncbi.nlm.nih.gov/25866929/)]
74. Sauro J, Lewis J. Correlations among prototypical usability metrics: evidence for the construct of usability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2009 Presented at: CHI '09: CHI Conference on Human Factors in Computing Systems; April 4 - 9, 2009; Boston MA USA p. 1609-1618. [doi: [10.1145/1518701.1518947](https://doi.org/10.1145/1518701.1518947)]
75. Copeland L, McNamara R, Kelson M, Simpson S. Mechanisms of change within motivational interviewing in relation to health behaviors outcomes: a systematic review. Patient Educ Couns 2015 Apr;98(4):401-411. [doi: [10.1016/j.pec.2014.11.022](https://doi.org/10.1016/j.pec.2014.11.022)] [Medline: [25535015](https://pubmed.ncbi.nlm.nih.gov/25535015/)]

Abbreviations

GBM: gay, bisexual, and other men who have sex with men
MI: motivational interviewing
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection
SUS: System Usability Scale
uMARS: Mobile App Rating Scale: User Version
UX: user experience

Edited by G Eysenbach; submitted 06.07.21; peer-reviewed by S John, C Klein, M Mulaudzi; comments to author 17.08.21; revised version received 08.09.21; accepted 26.09.21; published 16.11.21.

Please cite as:

Wray TB, Chan PA, Kahler CW

Game Plan, a Web-Based Intervention to Improve Adherence and Persistence to HIV Pre-exposure Prophylaxis and Reduce Heavy Drinking in Gay, Bisexual, and Other Men Who Have Sex With Men: Usability and User Experience Testing

JMIR Form Res 2021;5(11):e31843

URL: <https://formative.jmir.org/2021/11/e31843>

doi: [10.2196/31843](https://doi.org/10.2196/31843)

PMID: [34783662](https://pubmed.ncbi.nlm.nih.gov/34783662/)

©Tyler B Wray, Philip A Chan, Christopher W Kahler. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 16.11.2021. 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 4 Youth By Youth mHealth Photo Verification App for HIV Self-testing in Nigeria: Qualitative Analysis of User Experiences

David Oladele¹, MBBS, MSc; Juliet Iwelunmor², PhD; Titilola Gbajabiamila¹, MBBS, MSc; Chisom Obiezu-Umeh², MPH; Jane Ogoamaka Okwuzu¹, PhD; Ucheoma Nwaozuru², MS, PhD; Adesola Zaidat Musa¹, PhD; Ifeoma Idigbe¹, MSc; Kadija Tahlil³, MPH; Weiming Tang⁴, MD, PhD; Donaldson F Conserve⁵, PhD; Nora E Rosenberg³, PhD; Agatha N David¹, MBBS; Joseph Tucker^{6,7}, MD, PhD; Oliver Ezechi¹, MD, PhD

¹Clinical Sciences Department, Nigerian Institute of Medical Research, Lagos, Nigeria

²Department of Behavioral Science and Health Education, Saint Louis University, Saint Louis, MO, United States

³Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

⁴Dermatology Hospital, Southern Medical University, Guangzhou, China

⁵Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

⁶Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

⁷Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

Corresponding Author:

Juliet Iwelunmor, PhD

Department of Behavioral Science and Health Education

Saint Louis University

3545 Lafayette Avenue

Saint Louis, MO, 63104

United States

Phone: 1 3149773280

Email: juliet.iwelunmor@slu.edu

Abstract

Background: Despite the global expansion of HIV self-testing (HIVST), many research studies still rely on self-reported outcomes. New HIVST verification methods are needed, especially in resource-limited settings.

Objective: This study aims to evaluate the user experience of a mobile health (mHealth) app to enhance HIVST result reporting and verification.

Methods: Semistructured, in-depth interviews were used to evaluate the user experience of the 4 Youth By Youth mHealth photo verification app for HIVST. We used a think-aloud approach, and participants performed usability tasks and completed a qualitative exit interview. The app included HIV educational resources, step-by-step video instructions for performing HIVST, a 20-minute timer, a guide on interpreting results with linkages to care, an offline version, and a photo verification system. Demographic characteristics were reported by using descriptive statistics. Qualitative data were analyzed by using thematic analysis.

Results: A total of 19 users—12 women and 7 men—with a mean age of 22 years, participated in the study. The users completed the usability tasks and successfully uploaded a photo of their test results by using the app without assistance. Four main themes were identified in the data. First, in terms of user-friendly design, the participants noted the user-friendly features of the offline version and the app's low data use. However, some wanted the app to work in the background when using their mobile phone, and the font used should be more youth friendly. Second, in terms of ease of use, participants remarked that the app's self-explanatory nature and instructions that guided them on how to use the app enhanced its use. Third, in terms of a user's privacy, many participants reinforced the importance of privacy settings and tools that protect confidentiality among users. Finally, in terms of linkage to care, participants noted that the app's *linkage to care* features were useful, particularly in relation to referrals to trained counselors upon the completion of the test. All the participants noted that the app provided a convenient and private means of verifying the HIV test results.

Conclusions: Our findings demonstrated the importance of engaging end users in the development phase of health technology innovations that serve youth. Clinical trials are needed to determine the efficacy of using an mHealth app to verify HIVST results among young people.

(*JMIR Form Res* 2021;5(11):e25824) doi:[10.2196/25824](https://doi.org/10.2196/25824)

KEYWORDS

HIV self-testing; adolescents; young people; photo verification; mobile app; Nigeria

Introduction

Meeting the global 95-95-95 targets for HIV prevention among youth in sub-Saharan Africa requires innovation [1]. For example, in Nigeria, according to the 2019 National AIDS Indicator and Impact Survey report, there were 275,604 young people aged between 15 and 24 years living with HIV [2]. HIV self-testing (HIVST) is an innovation that has the potential to reach young people who may not otherwise test [3,4]. In this process, a person collects his or her own specimen (oral fluid or blood) and then performs a test and interprets the result, often in a private setting, either alone or with someone he or she trusts [5]. HIVST has been found to be an acceptable and effective strategy to increase HIV testing among marginalized populations [6,7], and the World Health Organization recommends HIVST as an approach to delivering HIV testing services to achieve the first 90 of the 90-90-90 strategy [5,8]. The Nigerian Federal Ministry of Health and the National Agency for Control of AIDS in Nigeria also recommended HIVST as a strategy to increase HIV testing uptake [9]. However, there are challenges with self-reporting of HIVST results because some people do not routinely self-report for a variety of reasons, and there is a need to enhance linkage to care following self-testing [8].

Understanding how best to reach young people in Nigeria with HIVST requires knowing whether they find HIVST feasible or acceptable. One formative study in Nigeria provided preliminary data on the willingness to engage in HIVST [10]. Moreover, in a cross-sectional survey of 157 individuals, 54.8% of participants expressed willingness to self-test, with some suggesting that HIVST will help avoid fear and stigma that may arise from the use of health facilities [11]. Similarly, in a pilot study among 257 men who have sex with men and key opinion leaders, the majority (98%) reported that it was easy to perform HIV testing with HIVST kits. Common reasons for liking HIVST were ease of use (87.3%), confidentiality and privacy (82.1%), convenience (74.1%), and absence of needle pricks (64.9%) [12]. HIVST kits became available in Nigeria in 2019. To date, only 2 HIVST kits have been introduced in Nigeria, which include the blood-based Alere HIV Combo rapid test kit and OraQuick self-test kit (saliva-based test), which is the only World Health Organization prequalified HIVST test in the country [13]. Despite the high rates of acceptability and willingness to use HIVST, we are not aware of any studies that have assessed how to improve access and reach young people with HIVST in Nigeria.

The need for this improved access for young people to HIVST kit in Nigeria prompted our youth-participatory research to increase HIVST among at-risk youth [13]. According to the Demographic and Health Survey in Nigeria 2018, 87.9% of

Nigerian households own mobile phones [14]. It has also been documented that expanding access to mobile health (mHealth) technology via the use of mobile devices has increased the opportunities to deliver health services to improve health outcomes [15-17]. Building on this infrastructure, mobile apps have the potential to (1) reach the largest number of young Nigerians with HIV who remain undiagnosed, (2) create demand for HIV prevention and increase HIV test coverage, and (3) facilitate linkage to treatment for young people who test HIV-positive and provide tailored prevention strategies for those who test HIV-negative.

In response, we partnered with Co-Creation Hub, Nigeria's first innovation laboratory and preincubation space for social tech ventures, to cocreate an HIVST mobile app targeting young Nigerians. The aim of the mobile app is to improve the verification of HIVST results and provide linkage to youth-friendly health services to young people who test for HIV. To accommodate the preferences and priorities of Nigerian youth, we engaged youth in the design and development of mobile HIVST apps. In this study, we described the development of the app and assessed the user experience. The findings will inform the implementation of mHealth HIVST services tailored to the needs of young people in Nigeria.

Methods

Development of the 4 Youth By Youth HIVST App

The Innovative Tools to Expand HIVST collaboration between the Nigerian Institute of Medical Research, Saint Louis University, New York University, and the University of North Carolina at Chapel Hill, known locally as 4 Youth By Youth (4YBY), engaged a team of developers at the Co-Creation Hub to develop the 4YBY photo verification mHealth app. The 4YBY photo verification mHealth app aims to promote HIVST by guiding users (youth) through the proper use of the HIVST kit and linking them to care and anonymous counseling. The mobile app was built with React Native to support the Android operating system from version 4.4 (Kitkat) upward. The development and design process were based on the Human-Centered Design framework. The design team worked together to solve 2 main challenges in creating the app: (1) how might we connect young people to health care and counseling when self-testing for HIV? (2) How can data on HIVST be collected and analyzed?

Some of the Key Activities in the Design Process

First, interviews were held with project key stakeholders, including potential users, with the goal to understand the facilitators and constraints or major concerns with HIVST. Second, there were immersion and shadowing activities

with potential users in which designers sought to understand and gain real-world experience with HIVST to understand end users' perception and preferences of HIVST. The goal of this process is to identify and provide value to young people who will use the product. Third, user interface prototyping and content creation was done. The design team worked with researchers from the Nigerian Institute of Medical Research to create content for the different sections of the mobile app. Given that the vast majority of people perceive images faster than words, we used illustrations supported by text to communicate important concepts throughout the design process.

Following the design phase, the front end and back end of the app were developed. A front-end section of the app, consisting of a visual representation of the app and a back-end section, also known as the server side of an app, including communication between the database and the app, were developed. After the development of the first version of the app, alpha and beta testing was performed. α testing was performed to identify bugs before releasing the app to end users, whereas beta testing was performed by potential end users of the app (young people aged 14-24 years) to determine the usability of the app.

The Key Features of the App

The key features of the app include the following:

1. How to use the HIVST test kit: It is a combination of illustrations and written step-by-step tutorials to help users perform their tests successfully (Figures 1-3).
2. Timer and reminder: Timers and reminders are used to help end users check for test results exactly 20 minutes after sample collection.
3. Photo upload and verification: Photo upload of the test kit at the expiration of the 20 minutes for the verification of test results. The purpose of this feature is to maintain data integrity and provide photo evidence of the HIVST result performed.
4. Linkage to care: This describes the process of linking the tester to care through access to toll-free calls for postcounseling after the self-test.

Other features were as follows:

1. HIV risk assessment quiz: It is a series of questions to help assess the tester's risk of contracting HIV and other sexually transmitted infections.
2. HIV facts and information: Educational content on HIV prevention, care, myths, and facts are included in the app that users are to review during the 20-minute period allotted for waiting for HIV test results.
3. Administrative dashboard: It is a web-based dashboard to access all data received through the mobile app. The dashboard will also allow counselors to reach out to users who have requested counseling.

Figure 1. Basic information on HIV in the app.

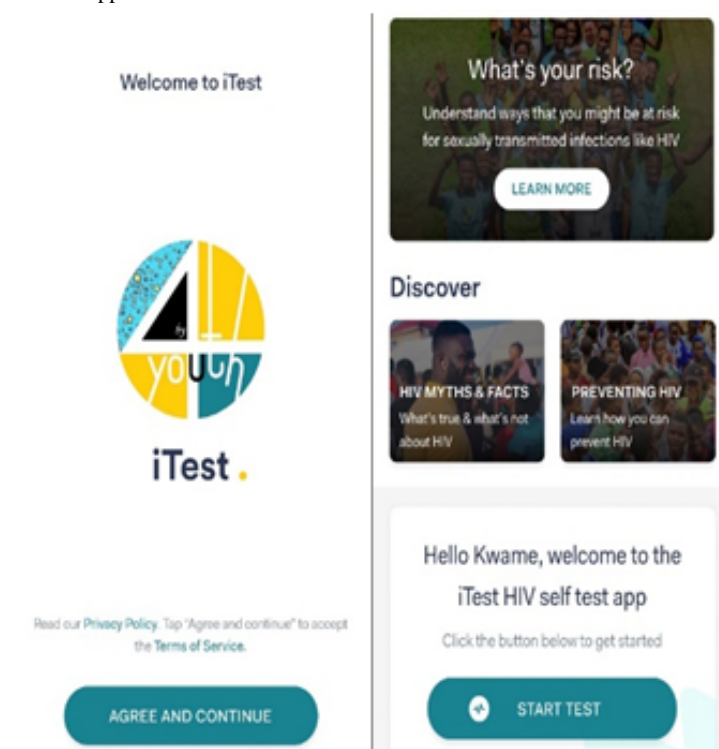


Figure 2. Screenshot of a video of the OraQuick swab.



Figure 3. Uploaded picture of an HIV test result on the app.



Study Design

A qualitative research design approach was used to evaluate the usability of the 4YBY mobile phone app by answering the following questions: To what extent do users achieve the desired goal of using the app (effectiveness)? Does the effort put into the use of the app allow the performance of a task in a timely manner (efficiency)? What is the perceived user's experience of the app? Participants used a think-aloud approach while performing usability tasks in three rounds of usability testing and subsequently completed semistructured one-on-one interviews.

Participant Recruitment

Participants were recruited through advertisements and word of mouth at university campuses and youth centers in Yaba, Lagos, Nigeria. Participants' recruitment for the study was guided by previous mHealth app studies that showed that preliminary usability of apps can be detected by a sample of 5 to 10 individuals [18,19]. Inclusion criteria included being aged between 14 and 24 years, fluency in English, the ability to read, owning an Android electronic device, and the ability to provide informed consent for the study. Data collection was performed by researchers who were part of the study team with requisite training in good clinical practice and ethical conduct of research.

Procedures

During each of the beta tests (Multimedia Appendix 1), participants performed HIVST using the OraQuick by OraSure testing kit. Tests were conducted by each participant swabbing their gum and then placing it in a processing fluid and waiting for 20 minutes for the result. Each participant had a prototype of the 4YBY app installed on their Android app and subsequently carried out the HIVST by following the directive provided by the app. The app also engaged the participants with basic information on HIV during the waiting time before the test completion, and they also participated in a HIV risk self-assessment test.

Data Collection

Following informed consent, participants were involved in usability testing via the think-aloud approach [20]. The think-aloud approach required users to continuously verbalize their thoughts about their underlying thinking behind their interactions while using the mobile app by freely expressing what they were doing and why and stating when they encountered any problem in the process. This provided an opportunity for participants to test the app's functionalities and provide opinions, comments, and concerns to improve on the app. They also discussed how they felt about the use of the app, which was audiotaped.

Immediately following usability testing, participants were asked a series of questions to assess their experience with the usability of the app (Multimedia Appendix 2). The objective of the phase was to obtain the participants' immediate interpretation of the app design and functionality and facilitate the elaboration of usability issues and increasing insight and design suggestions [21]. A total of 3 beta tests using this methodology were performed over 3 months. After each of the beta tests, the design team used the user experience feedback to improve the development of the app by correcting all glitches noted and modifying as necessary to meet users.

Data Analysis

The demographic characteristics of the users were reported as descriptive statistics. Each interview was audio recorded, subsequently transcribed verbatim, and manually coded. Field notes were also included in the analysis. Inductive thematic analysis, which describes the process of coding data without trying to fit the data into a pre-existing coding frame or the researcher's analytic preconceptions, was used for this study. This process ensured that thematic analysis was completely data driven [22]. The Braun and Clark [22,23] guidelines were used by 2 research team members with experience in qualitative research, which included the following steps: (1) repeated reading of the transcripts to become familiar with the data; (2) generating initial codes relevant to the research questions (effectiveness, efficiency, and usefulness); (3) organizing the codes; (4) arranging the subthemes into overarching themes; and (5) defining and naming themes. Disagreements around themes were resolved by an in-person research team discussion until consensus was reached, and a final theme was agreed upon.

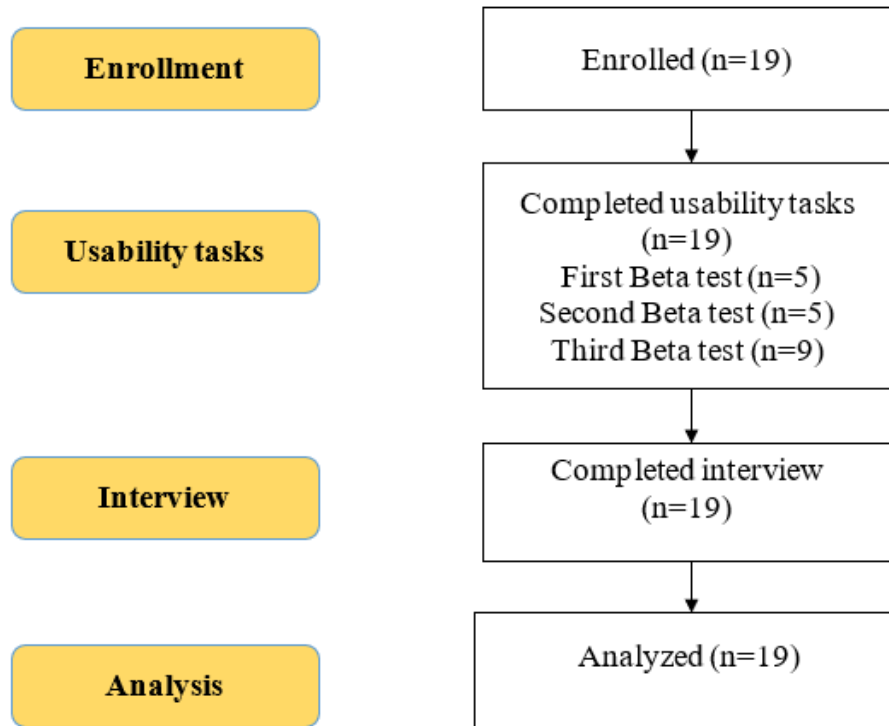
Ethical Consideration

Ethical approval for the Innovative Tools to Expand HIVST study was obtained from the institutional review board of the Nigerian Institute of Medical Research and Saint Louis University. Informed consent was obtained from each participant before the commencement of the usability beta testing study.

Results

Participant Characteristics and Usability Findings

A total of 19 youth participated in the usability analysis of the mHealth app for HIVST (Figure 4). The mean age of the participants was 22 years (SD 2.4), and the age range was 14-25 years. There were 12 females and 7 males and 68% (13/19) of the participants were students. The majority (18/19, 95%) of the participants completed all the usability tasks on the first attempt. On average, the usability test lasted between 25 and 30 minutes. All the participants successfully uploaded a photo of their test results using the app. Figure 4 illustrates participants' progression through each of the study phases.

Figure 4. Study flow diagram.

User Experiences

Thematic analysis of the user comments from the think-aloud interviews from the 3 beta usability tests yielded findings on how the app was used, including the relation to test outcomes

(ie, positive, negative, and invalid) and strengths and weaknesses of the prototype app and suggested improvements. The following four themes were identified (Table 1): (1) the design of the app, (2) the ease of use of the app, (3) user privacy reinforcement, and (4) linkage to care.

Table 1. Summary of finding from one-on-one interviews on usability of the 4 Youth By Youth app.

Variables considered	Beta test 1 (August 1, 2019)	Beta test 2 (September 24, 2019)	Beta test 3 (September 26, 2019)
Design of the app	<ul style="list-style-type: none"> “The App could be more attractive, with regards to color” 	<ul style="list-style-type: none"> “The App should be more colorful/attractive” “There is a need to improve on the resolution of the videos” 	<ul style="list-style-type: none"> “The App background should be more attractive” “There is a need to use more phrases that will catch Nigerian youth attention”
Ease of use of the app	<ul style="list-style-type: none"> “It is easy to use the App. The picture/videos were helpful even for a layman” “It is simple and straight forward” “Basic HIV information on the App was useful” 	<ul style="list-style-type: none"> “The information about HIV on the App was quite educative” “The instruction on HIV self-testing on the App is adequate to safely conduct the test” “The App instruction on how to setup the test need a little more work” “I would prefer if the language of the App could also be pidgin English” “The App contain all the information I will need to do HIV self-testing on my own” 	<ul style="list-style-type: none"> “The App is explicit, self-explanatory” “The video made it easy to understand the test process” “I wonder if the App could be used offline” “It was easy to take the photo evidence of my test” “The explanation of the test result is adequate to educate user on their test result, its significance and next steps to be taken”
App use and privacy	<ul style="list-style-type: none"> “It does encourage privacy” 	<ul style="list-style-type: none"> “The App will promote privacy because many young people want to keep their health issues private” “The App is convenient especially for someone who do not want to go to health center or who live far away from health center” 	<ul style="list-style-type: none"> “Am happy that no one can have access to my result by looking at the App on my phone”
Linkage to care after checking results	<ul style="list-style-type: none"> “After reading my result, I will prefer to call the health facility with the number provided” 	<ul style="list-style-type: none"> “The feature on the App to request a call or call a youth counselor is great, though I would prefer to make the call myself after seeing my result” 	<ul style="list-style-type: none"> “Yes I will request a call from the youth counselor if my result is positive”
Recommend the app to peers	<ul style="list-style-type: none"> N/A^a 	<ul style="list-style-type: none"> N/A 	<ul style="list-style-type: none"> “I will recommend the App to my friends because it is simple and easy to use”
Timer countdown to result	<ul style="list-style-type: none"> “The time of waiting to read the result is too long.” “I will prefer an alert/buzz when the 20 minutes waiting time elapses” 	<ul style="list-style-type: none"> N/A 	<ul style="list-style-type: none"> “The buzz at the end of 20 minutes is important. I also like the fact that the App could function at the background while I wait for my result”

^aN/A: not applicable.

Design of the App

Participants frequently commented on the need for the app to have a more user-friendly design, with interactive images, an appealing color scheme, and features that are easily accessible. The participants involved in the first beta testing felt that a mobile app intended for use by youth should be colorful and attractive to ensure adequate coverage and acceptability among this age group. They further explained that the likelihood of a young person downloading the app from the Google Play store depends largely on how they perceived the app:

The App could be more attractive with regards to its color. [First round beta test participant]

During the second and third beta tests, participants also wanted an improvement in the appearance of the logo of the app and its background theme. Other concerns raised during the second

beta test are the need to improve the resolution of the video that describes the process of HIVST in the app:

There is a need to improve the resolution of the videos. [Second beta test participant]

A participant in the third beta test also expressed a preference that the phrases used in the app should be the type that will easily catch the attention of Nigerian youths who are the end users of the app. He would prefer developers to use some locally trending phrases on social media:

There is a need to use more phrases that will catch Nigerian youth attention. [Third beta test participant]

During the first beta test, some participants noted that it would be desirable to have the app work in the background while he could be engaged in the use of the mobile phone for other things. The feature was corrected before the third beta testing, and

participants at the third beta testing were impressed with this feature:

The time of waiting to read the result is too long. I will prefer an alert/buzz when the 20 minutes waiting time elapses. [First beta test participant]

The buzz at the end of 20 minutes is important. I also like the fact that the App could function at the background while I wait for my result. [Third beta test participant]

Ease of Use of the App

The ease of use of the app assesses the user experience of the app without direct guidance from the developers. Participants in each of the beta tests overall found the app easy to use and did not report that anything went wrong during the use of the app. Some of the other observations made on the ease of use include the need to adjust the app directive on how to set up the test and if the app could be used offline without internet access:

It is easy to use the App. The picture/videos were helpful even for a layman, It is simple and straight-forward. [First beta test participant]

The instruction on HIV self-testing on the App is adequate to safely conduct the test. The app instruction on how to set up the test needs a little more work. I would prefer it if the language of the App could also be Pidgin English. The app contains all the information I will need to do HIV self-testing on my own. [Second beta test participant]

The App is explicit, self-explanatory. The video made it easy to understand the test process. I wonder if the App could be used offline. It was easy to take the photo evidence of my test. The explanation of the test result is adequate to educate users on their test result, its significance, and the next steps to be taken. [Third beta test participant]

User Privacy Reinforcement

HIV stigma is a key pertinent issue with HIV testing. To ensure that all users can test anonymously, aliases, rather than real names, were used. The developer also implemented an auto-logout feature for further security. However, in think-aloud exercises, one of the major concerns shared among participants was the issue of privacy. The participants would not want their personal information to be shared in any form, and the app ensured anonymity to address this. The participants during the 3 phases of the beta test noted that the app further reinforced the practice of privacy in getting tested for HIV:

The App will promote privacy because many young people want to keep their health issues private. The App is convenient especially for someone who does not want to go to the health center or who lives far away from the health center. [Second beta test participant]

Am happy that no one can have access to my result by looking at the App on my phone. [Third beta test participant]

Linkage to Services

One of the features of the 4YBY photo verification app is that the participants can be linked to care after conducting an HIV test. This is usually done by the user calling a counselor to discuss the HIV test result or requesting a call from a counselor by clicking on a radio button in the app. The users were pleased that this feature was included in the app, and they would either call or request a call to be linked to care after performing the HIV self-test:

After reading my result, I will prefer to call the health facility with the number provided. [First beta test participant]

The feature on the App to request a call or call a youth counselor is great, though I would prefer to make the call myself after seeing my result [Second beta test participant]

Yes I will request a call from the youth counselor if my result is positive. [Third beta test participant]

Discussion

Principal Findings

We developed a photo verification mHealth app to promote HIVST among young people in Nigeria. To our knowledge, this is one of the first mHealth apps that has been developed to promote HIVST and linkage to care in sub-Saharan Africa [24]. Our usability findings add evidence to the limited data on young people's experiences and perspectives on mHealth apps in sub-Saharan Africa. We identified 4 main themes (eg, app design, ease of use, user privacy reinforcement, and linkage to care) that may influence the adoption and continued use of the apps for HIVST.

This study corroborates the findings of limited studies in sub-Saharan Africa, including how mHealth apps for HIVST may overcome the experience of stigma and structural barriers of clinic-based testing by providing privacy and convenience [25]. Even in middle- to high-income countries, with a substantial body of research on mHealth apps and HIVST, findings suggest that the design of these apps, along with their ease of use, privacy, and linkage to services, provides an opportunity to disseminate information about the accuracy, safety, and appropriate use of this technology to promote HIVST [26].

Our iterative cocreation process led to the development of an mHealth HIVST app suited to the needs of young people in Nigeria. In addition, the iterative process and feedback from beta testing led to the modification of the app to address issues raised during the development phase, including the desire to have an app with minimal internet data demand, features to encourage users' anonymity, and step-by-step video description of the process of conducting the HIVST. Previous studies on mHealth app development have documented the importance of involving end users in the development of mHealth apps [27]. Through their own experiences and by showing participants how to use our app, the majority found the app to be informative, accessible, and engaging. Think-aloud activities identified not

only facilitating features that may potentially enhance uptake but also factors that may hinder its use.

Consistent with other mHealth app literature [28,29], one primary recommendation was for the app to reinforce privacy concerns. This is particularly salient for young people, given their limited access to youth-friendly health centers and barriers associated with navigating and appearing at clinics for HIV prevention services that mostly cater to adults [30,31]. Protecting the privacy of young people accessing the app will not only maintain youth support but also enhance trust. Providing linkage to additional services via the app also provides an opportunity to further enhance youth engagement with HIV prevention, treatment, and care.

Participants noted that the app's linkage platform can enable HIV services to reach young people at risk for or living with HIV, who may not otherwise use health care services. In addition, to the linkage to services, participants reported that the use of multimedia text, short videos, and graphics within the app may not only help youth to remain HIV-uninfected but also enhance their ability to stay healthy and thrive if they are already living with HIV. Overall, these findings provide insight into young people's concerns and strategies to enhance the uptake of our mHealth app for HIVST promotion.

This study had some limitations. First, this was a small sample of youth participants recruited from Lagos, Nigeria, thereby making it difficult to reach generalizations from this data source. Second, access to smartphones and mHealth apps may have influenced social desirability bias with the decision to participate in the study. In addition, the flexibility of thematic analysis can lead to inconsistency and a lack of coherence with theme development. To address authors' bias, the coding of the transcript was independently performed by 2 researchers before theme development. Finally, given that this is a usability study, it is not possible to conclude the potential efficacy of the app in promoting HIVST uptake among young people in Nigeria.

The results of this study have implications for interventions targeting HIVST uptake among young people in Nigeria. It also has important implications for using technology to disseminate HIV prevention information to youth populations in Nigeria.

The 2019 Nigerian operational guidelines for HIVST seek to increase individuals' autonomy, decentralize services, and create demand for HIV testing, especially for underserved populations, including young people. Given the new operational guidelines that seek to decentralize HIV testing, verification of self-reported outcomes for HIVST is of utmost importance. Bearing in mind that there are no available photo verification modalities for HIVST for young people in Nigeria, this study provides evidence on the usability of the 4YBY mHealth photo verification app and its potential to promote uptake of HIV testing among young people. These findings lay the foundation for future studies to assess the feasibility, reach, and effectiveness of the mHealth app to promote the uptake of HIV testing and other preventive services among young people in Nigeria. Through this formative work, we also obtained contextual recommendations from the participants to improve the design of the app and reinforce privacy to enhance acceptability among young people in Nigeria. The feedback has practical implications for mHealth technologies to seek end users' feedback to ensure acceptability and appropriateness among target populations. The 4YBY mHealth app may provide a model for supporting young people to access other sexual and reproductive health services, given the app's built-in function to link users to available youth-friendly health facilities. The findings have implications for practice in Nigeria and can inform the integration of mHealth technologies within youth-friendly health facilities to promote the uptake of preventive sexual health services.

Conclusions

This study demonstrated the usability of an mHealth app to promote HIVST uptake among young people in Nigeria. Findings from this study illustrated that the design and ease of use of the app, along with privacy and linkage to additional services, may enhance the reach and acceptability of HIVST services among young people who are traditionally underserved in traditional HIV research conducted not only in Nigeria but also in sub-Saharan Africa. Future clinical trials are needed to determine the efficacy of using an mHealth app to verify HIVST results among young people.

Acknowledgments

The authors would like to thank the Nigerian Institute of Medical Research and Co-Creation Hub Nigeria. The research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development under award UG3HD096929. JT was supported by the National Institute of Allergy and Infectious Diseases under award K24AI143471. The funders had no role in the design of the study; the collection, analysis, and interpretation of data; or writing of the manuscript.

Authors' Contributions

DO, JI, and OE conceived the ideas for the study. DO, UN, and COU drafted the manuscript. DO, UN, COU, KT, II, and TG performed data acquisition and data analysis. AZM, TG, AND, WT, DFC, and NER reviewed the drafts and provided written feedback. JT, JI, and OE edited the paper for the critical content. All authors contributed substantially to the preparation of this manuscript. All authors have read and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

App process flow (usability task flowchart).

[\[PDF File \(Adobe PDF File\), 264 KB - formative_v5i11e25824_app1.pdf \]](#)

Multimedia Appendix 2

Usability questionnaire.

[\[DOCX File , 14 KB - formative_v5i11e25824_app2.docx \]](#)

References

1. Communities at The Centre: Defending Rights, Breaking Barriers, Reaching People With HIV Services. reliefweb. 2019. URL: <https://reliefweb.int/report/world/communities-centre-defending-rights-breaking-barriers-reaching-people-hiv-services> [accessed 2021-08-28]
2. NAIIS national summary sheet. National Agency for Control of AIDS. URL: <https://naca.gov.ng/naiis-national-summary-sheet/> [accessed 2021-08-28]
3. Indravudh P, Sibanda EL, d'Elbée M, Kumwenda MK, Ringwald B, Maringwa G, et al. 'I will choose when to test, where I want to test': investigating young people's preferences for HIV self-testing in Malawi and Zimbabwe. *AIDS* 2017 Jul 01;31 Suppl 3(Suppl 3):S203-S212 [FREE Full text] [doi: [10.1097/QAD.0000000000001516](https://doi.org/10.1097/QAD.0000000000001516)] [Medline: [28665878](https://pubmed.ncbi.nlm.nih.gov/28665878/)]
4. Hatzold K, Gudukeya S, Mutseta MN, Chilongosi R, Nalubamba M, Nkhoma C, et al. HIV self-testing: breaking the barriers to uptake of testing among men and adolescents in sub-Saharan Africa, experiences from STAR demonstration projects in Malawi, Zambia and Zimbabwe. *J Int AIDS Soc* 2019 Mar;22 Suppl 1(Suppl 1):e25244 [FREE Full text] [doi: [10.1002/jia2.25244](https://doi.org/10.1002/jia2.25244)] [Medline: [30907505](https://pubmed.ncbi.nlm.nih.gov/30907505/)]
5. Guidelines on HIV Self-testing and Partner Notification: Supplement to Consolidated Guidelines on HIV Testing Services. Geneva: World Health Organization; Dec 2016.
6. Krause J, Subklew-Sehume F, Kenyon C, Colebunders R. Acceptability of HIV self-testing: a systematic literature review. *BMC Public Health* 2013 Aug 08;13:735 [FREE Full text] [doi: [10.1186/1471-2458-13-735](https://doi.org/10.1186/1471-2458-13-735)] [Medline: [23924387](https://pubmed.ncbi.nlm.nih.gov/23924387/)]
7. Choko AT, MacPherson P, Webb EL, Willey BA, Feasy H, Sambakunsi R, et al. Uptake, accuracy, safety, and linkage into care over two years of promoting annual self-testing for HIV in Blantyre, Malawi: a community-based prospective study. *PLoS Med* 2015 Sep 8;12(9):e1001873 [FREE Full text] [doi: [10.1371/journal.pmed.1001873](https://doi.org/10.1371/journal.pmed.1001873)] [Medline: [26348035](https://pubmed.ncbi.nlm.nih.gov/26348035/)]
8. WHO recommends HIV self-testing ? Evidence update and considerations for success. World Health Organization. 2019. URL: <https://www.who.int/publications/i/item/WHO-CDS-HIV-19.36> [accessed 2021-08-28]
9. Federal MOH, Nigeria. HIV self testing : operational guidelines for the delivery of HIV self-testing in Nigeria. - 2018:2018. [doi: [10.14293/s2199-1006.1.sor-med.cl8okwo.v1](https://doi.org/10.14293/s2199-1006.1.sor-med.cl8okwo.v1)]
10. Nwaozuru U, Iwelunmor J, Ong JJ, Salah S, Obiezu-Umeh C, Ezechi O, et al. Preferences for HIV testing services among young people in Nigeria. *BMC Health Serv Res* 2019 Dec 27;19(1):1003 [FREE Full text] [doi: [10.1186/s12913-019-4847-x](https://doi.org/10.1186/s12913-019-4847-x)] [Medline: [31881959](https://pubmed.ncbi.nlm.nih.gov/31881959/)]
11. Brown B, Folayan MO, Imosili A, Durueke F, Amuamuziam A. HIV self-testing in Nigeria: public opinions and perspectives. *Glob Public Health* 2015;10(3):354-365. [doi: [10.1080/17441692.2014.947303](https://doi.org/10.1080/17441692.2014.947303)] [Medline: [25186234](https://pubmed.ncbi.nlm.nih.gov/25186234/)]
12. Tun W, Vu L, Dirisu O, Sekoni A, Shoyemi E, Njab J, et al. Uptake of HIV self-testing and linkage to treatment among men who have sex with men (MSM) in Nigeria: a pilot programme using key opinion leaders to reach MSM. *J Int AIDS Soc* 2018 Jul;21 Suppl 5:e25124 [FREE Full text] [doi: [10.1002/jia2.25124](https://doi.org/10.1002/jia2.25124)] [Medline: [30033680](https://pubmed.ncbi.nlm.nih.gov/30033680/)]
13. Obiezu-Umeh C, Gbajabiamila T, Ezechi O, Nwaozuru U, Ong JJ, Idigbe I, et al. Young people's preferences for HIV self-testing services in Nigeria: a qualitative analysis. *BMC Public Health* 2021 Jan 07;21(1):67 [FREE Full text] [doi: [10.1186/s12889-020-10072-1](https://doi.org/10.1186/s12889-020-10072-1)] [Medline: [33413246](https://pubmed.ncbi.nlm.nih.gov/33413246/)]
14. Iwelunmor J, Ezechi O, Obiezu-Umeh C, Gbaja-Biamila T, Nwaozuru U, Oladele D, et al. The 4 youth by youth HIV self-testing crowdsourcing contest: a qualitative evaluation. *PLoS One* 2020 May 29;15(5):e0233698 [FREE Full text] [doi: [10.1371/journal.pone.0233698](https://doi.org/10.1371/journal.pone.0233698)] [Medline: [32469971](https://pubmed.ncbi.nlm.nih.gov/32469971/)]
15. Nigeria Demographic and Health Survey 2018. National Population Commission. 2019. URL: <https://dhsprogram.com/pubs/pdf/FR359/FR359.pdf> [accessed 2021-08-28]
16. The State of ICT In Nigeria 2018. Research ICT Africa. URL: <https://researchictafrica.net/after-access-nigeria-state-of-ict-2017/> [accessed 2021-08-28]
17. Krah EF, de Kruijff JG. Exploring the ambivalent evidence base of mobile health (mHealth): a systematic literature review on the use of mobile phones for the improvement of community health in Africa. *Digit Health* 2016 Nov 24;2:2055207616679264 [FREE Full text] [doi: [10.1177/2055207616679264](https://doi.org/10.1177/2055207616679264)] [Medline: [29942576](https://pubmed.ncbi.nlm.nih.gov/29942576/)]
18. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. *JAMA Pediatr* 2017 May 01;171(5):461-469 [FREE Full text] [doi: [10.1001/jamapediatrics.2017.0042](https://doi.org/10.1001/jamapediatrics.2017.0042)] [Medline: [28319239](https://pubmed.ncbi.nlm.nih.gov/28319239/)]

19. Holtz BE, Murray KM, Hershey DD, Dunneback JK, Cotten SR, Holmstrom AJ, et al. Developing a patient-centered mHealth app: a tool for adolescents with type 1 diabetes and their parents. *JMIR Mhealth Uhealth* 2017 Apr 19;5(4):e53 [[FREE Full text](#)] [doi: [10.2196/mhealth.6654](https://doi.org/10.2196/mhealth.6654)] [Medline: [28428167](https://pubmed.ncbi.nlm.nih.gov/28428167/)]
20. Beatty AL, Magnusson SL, Fortney JC, Sayre GG, Whooley MA. VA FitHeart, a mobile app for cardiac rehabilitation: usability study. *JMIR Hum Factors* 2018 Jan 15;5(1):e3 [[FREE Full text](#)] [doi: [10.2196/humanfactors.8017](https://doi.org/10.2196/humanfactors.8017)] [Medline: [29335235](https://pubmed.ncbi.nlm.nih.gov/29335235/)]
21. The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes. London: Academic Press; 1994.
22. Følstad A, Hornbæk K. Work-domain knowledge in usability evaluation: experiences with cooperative usability testing. *J Syst Softw* 2010 Nov;83(11):2019-2030. [doi: [10.1016/j.jss.2010.02.026](https://doi.org/10.1016/j.jss.2010.02.026)]
23. Clarke V, Braun V, Hayfield N. Thematic analysis. In: *Qualitative Psychology: A Practical Guide to Research Methods*. London, UK: Sage Publication; 2015.
24. Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. In: *The SAGE Handbook of Qualitative Research in Psychology*. Thousand Oaks, CA: Sage Publication; 2017.
25. Tahlil KM, Ong JJ, Rosenberg NE, Tang W, Conserve DF, Nkengasong S, et al. Verification of HIV self-testing use and results: a global systematic review. *AIDS Patient Care STDS* 2020 Apr;34(4):147-156 [[FREE Full text](#)] [doi: [10.1089/apc.2019.0283](https://doi.org/10.1089/apc.2019.0283)] [Medline: [32324482](https://pubmed.ncbi.nlm.nih.gov/32324482/)]
26. 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](https://pubmed.ncbi.nlm.nih.gov/31738174/)]
27. 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/)]
28. Feldacker C, Holeman I, Murenje V, Xaba S, Korir M, Wambua B, et al. Usability and acceptability of a two-way texting intervention for post-operative follow-up for voluntary medical male circumcision in Zimbabwe. *PLoS One* 2020 Jun 16;15(6):e0233234 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0233234](https://doi.org/10.1371/journal.pone.0233234)] [Medline: [32544161](https://pubmed.ncbi.nlm.nih.gov/32544161/)]
29. Gkatzidou V, Hone K, Sutcliffe L, Gibbs J, Sadiq ST, Szczepura A, et al. User interface design for mobile-based sexual health interventions for young people: design recommendations from a qualitative study on an online Chlamydia clinical care pathway. *BMC Med Inform Decis Mak* 2015 Aug 26;15:72 [[FREE Full text](#)] [doi: [10.1186/s12911-015-0197-8](https://doi.org/10.1186/s12911-015-0197-8)] [Medline: [26307056](https://pubmed.ncbi.nlm.nih.gov/26307056/)]
30. Feroz A, Abrejo F, Ali SA, Nuruddin R, Saleem S. Using mobile phones to improve young people's sexual and reproductive health in low- and middle-income countries: a systematic review protocol to identify barriers, facilitators and reported interventions. *Syst Rev* 2019 May 18;8(1):117 [[FREE Full text](#)] [doi: [10.1186/s13643-019-1033-5](https://doi.org/10.1186/s13643-019-1033-5)] [Medline: [31103044](https://pubmed.ncbi.nlm.nih.gov/31103044/)]
31. Femi-Adebayo T, Kuyinu Y, Adejumo OA, Goodman O. Factors affecting utilization of youth friendly health services in Lagos State, Nigeria. *Int J Adolesc Med Health* 2017 Sep 15;31(2):/j/ijamh.2019.31.issue-2/ijamh-2016-0159/ijamh-2016-0159.xml. [doi: [10.1515/ijamh-2016-0159](https://doi.org/10.1515/ijamh-2016-0159)] [Medline: [28915109](https://pubmed.ncbi.nlm.nih.gov/28915109/)]

Abbreviations

4YBY: 4 Youth By Youth

HIVST: HIV self-testing

mHealth: mobile health

Edited by G Eysenbach; submitted 17.11.20; peer-reviewed by Y Zhang, A Hoskins, S Butler; comments to author 01.03.21; revised version received 31.05.21; accepted 01.08.21; published 17.11.21.

Please cite as:

Oladele D, Iwelunmor J, Gbajabiamila T, Obiezu-Umeh C, Okwuzu JO, Nwaozuru U, Musa AZ, Idigbe I, Tahlil K, Tang W, Conserve DF, Rosenberg NE, David AN, Tucker J, Ezechi O

The 4 Youth By Youth mHealth Photo Verification App for HIV Self-testing in Nigeria: Qualitative Analysis of User Experiences
JMIR Form Res 2021;5(11):e25824

URL: <https://formative.jmir.org/2021/11/e25824>

doi: [10.2196/25824](https://doi.org/10.2196/25824)

PMID: [34787579](https://pubmed.ncbi.nlm.nih.gov/34787579/)

©David Oladele, Juliet Iwelunmor, Titilola Gbajabiamila, Chisom Obiezu-Umeh, Jane Ogoamaka Okwuzu, Ucheoma Nwaozuru, Adesola Zaidat Musa, Ifeoma Idigbe, Kadija Tahlil, Weiming Tang, Donaldson F Conserve, Nora E Rosenberg, Agatha N David,

Joseph Tucker, Oliver Ezechi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.11.2021. 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 Two-Minute Walking Test With a Smartphone App for Persons With Multiple Sclerosis: Validation Study

Pim van Oirschot¹, PhD; Marco Heerings^{2,3}, MANP; Karine Wendrich⁴, MSc; Bram den Teuling¹, MSc; Frank Dorssers¹, MSc; René van Ee^{1,5}, MEng, MD; Marijn Bart Martens^{6,7}, PhD; Peter Joseph Jongen^{8,9}, PhD, MD

¹Orikami Digital Health Products, Nijmegen, Netherlands

²Dutch National Multiple Sclerosis Foundation, Rotterdam, Netherlands

³Radboud University Medical Center, Nijmegen, Netherlands

⁴Faculty of Science, Institute for Science in Society, Radboud University, Nijmegen, Netherlands

⁵Sint Maartenskliniek, Nijmegen, Netherlands

⁶Drug Target ID, Nijmegen, Netherlands

⁷NeuroDrug Research BV, Nijmegen, Netherlands

⁸Department of Community & Occupational Medicine, University Medical Centre Groningen, Groningen, Netherlands

⁹MS4 Research Institute, Nijmegen, Netherlands

Corresponding Author:

Pim van Oirschot, PhD

Orikami Digital Health Products

Ridderstraat 29

Nijmegen, 6511 TM

Netherlands

Phone: 31 24 301 0100

Email: pim@mssherpa.nl

Abstract

Background: Walking disturbances are a common dysfunction in persons with multiple sclerosis (MS). The 2-Minute Walking Test (2MWT) is widely used to quantify walking speed. We implemented a smartphone-based 2MWT (s2MWT) in MS sherpa, an app for persons with MS. When performing the s2MWT, users of the app are instructed to walk as fast as safely possible for 2 minutes in the open air, while the app records their movement and calculates the distance walked.

Objective: The aim of this study is to investigate the concurrent validity and test-retest reliability of the MS sherpa s2MWT.

Methods: We performed a validation study on 25 persons with relapsing-remitting MS and 79 healthy control (HC) participants. In the HC group, 21 participants were matched to the persons with MS based on age, gender, and education and these followed the same assessment schedule as the persons with MS (the *HC-matched* group), whereas 58 participants had a less intense assessment schedule to determine reference values (the *HC-normative* group). Intraclass correlation coefficients (ICCs) were determined between the distance measured by the s2MWT and the distance measured using distance markers on the pavement during these s2MWT assessments. ICCs were also determined for test-retest reliability and derived from 10 smartphone tests per study participant, with 3 days in between each test. We interviewed 7 study participants with MS regarding their experiences with the s2MWT.

Results: In total, 755 s2MWTs were completed. The adherence rate for the persons with MS and the participants in the HC-matched group was 92.4% (425/460). The calculated distance walked on the s2MWT was, on average, 8.43 m or 5% (SD 18.9 m or 11%) higher than the distance measured using distance markers (n=43). An ICC of 0.817 was found for the concurrent validity of the s2MWT in the combined analysis of persons with MS and HC participants. Average ICCs of 9 test-retest reliability analyses of the s2MWT for persons with MS and the participants in the HC-matched group were 0.648 (SD 0.150) and 0.600 (SD 0.090), respectively, whereas the average ICC of 2 test-retest reliability analyses of the s2MWT for the participants in the HC-normative group was 0.700 (SD 0.029). The interviewed study participants found the s2MWT easy to perform, but they also expressed that the test results can be confronting and that a pressure to reach a certain distance can be experienced.

Conclusions: The high correlation between s2MWT distance and the conventional 2MWT distance indicates a good concurrent validity. Similarly, high correlations underpin a good test-retest reliability of the s2MWT. We conclude that the s2MWT can be

used to measure the distance that the persons with MS walk in 2 minutes outdoors near their home, from which both clinical studies and clinical practice can benefit.

(*JMIR Form Res* 2021;5(11):e29128) doi:[10.2196/29128](https://doi.org/10.2196/29128)

KEYWORDS

multiple sclerosis; relapsing remitting; mobility; mobile phone; 2-Minute Walking Test

Introduction

Background

Multiple sclerosis (MS) is a common and as yet incurable chronic neurological condition [1]. It affects the central nervous system, involving demyelination and resulting in impairment of the nerve conduction. MS symptoms are different for virtually every person with MS, and the course of the disease is unpredictable. Symptoms of MS include, among others, fatigue, cognitive impairment, visual disturbances, sensory disturbances, and balance problems. For some persons with MS, the disease progresses slowly, while others experience years of rapidly increasing disability soon after the diagnosis. MS is most commonly diagnosed in people in their 20s and 30s, although it can develop at any age. MS affects 3 times more women than men.

Approximately 75% of the persons with MS experience a clinically significant walking disturbance, which poses a barrier to being physically active [2]. Mobility problems may arise from various factors, such as fatigue, decreasing muscle strength, spasticity, and ataxia and may result in musculoskeletal pain in the back, hips, legs, and arms. Pain in turn may further reduce walking ability. Walking limitations are a key component of disability in persons with MS. The Expanded Disability Status Scale (EDSS), which ranges from 0 to 10 and is universally used by health professionals to quantify physical disability in persons with MS, relies on walking as the main measure of disability [3]. Physical activity is highly relevant for patients because it has a large impact on employment and generally on the quality of life [4]. Mobility supports social participation and, in many cases, the ability to work, which is important to prevent persons with MS feeling isolated and depressed [5].

Several walking tests are available for persons with MS, including the Timed Up and Go Test, the Timed 25-Foot Walk (T25FW), and 6-Minute Walking Test (6MWT).

The Timed Up and Go test measures the time that a person takes to rise from a chair, walk 3 m, turn around, walk back to the chair, and sit down. The T25FW measures the time that a person takes to walk 25 feet, approximately 7.6 m. In the 6MWT, the distance walked in 6 minutes is measured. Bennett et al [6] showed that the distance walked on a 2-Minute Walking Test (2MWT), a 3 times shorter variant of the 6MWT, highly correlates with Timed Up and Go, T25FW, and 6MWT scores and that it also correlates with EDSS.

Walking assessments for persons with MS can currently be scheduled upon request of their health care professionals. Usually, such an assessment is done in the MS clinic or at the physiotherapist. Recent technological advances show promise that in the near future, walking tests might be performed in the

home environment of a person with MS [7-15]. This saves time and costs and makes it possible to schedule assessments more frequently. The wealth of data from regularly performed home assessments could also improve clinical decision-making because it provides health care providers with quantitative information about changes in their patients' walking speed.

MS sherpa (Orikami Digital Health Products) is a software used as a medical device and intended to support the monitoring of persons with MS to give patients and their health care professionals personalized insight into the presence and progress of MS-related symptoms and signs [16-20]. MS sherpa is a system consisting of a smartphone app (supported on Android and iOS) for data collection and data presentation, a cloud service for data storage, analysis algorithms, and a clinician or researcher dashboard for user management and data presentation. The product is commercially available. More information can be found on the MS sherpa website [20].

It is possible to do a smartphone 2MWT (s2MWT) with MS sherpa. Instructions in an explanatory text in the app include that users should walk outside as fast as safely possible while still not running or jogging using a walking aid if necessary, with their phone in their trouser pocket during the test. Once an accurate GPS location signal is found, the test can be started. After the start button is pressed, users can place their smartphone in their trouser pocket during a 5-second countdown. At the end of the countdown, the users should start walking, and they can stop when they feel a vibration and hear a sound exactly 2 minutes later. Then, the distance walked is calculated and the test result is displayed in the app. The s2MWT differs in concept from traditional 2MWTs, which are generally performed indoors by walking on a level surface between 2 lines with a known distance, with various methods (and accuracies) for determining the length of the final stretch.

Objective

The aim of this study is to investigate the concurrent validity and test-retest reliability of the s2MWT that is implemented in MS sherpa. The validation of the s2MWT was part of MS Self—a validation study during which participants performed self-monitoring assessments during 4 weeks with a precursor of MS sherpa, the *Mijn Kwik* (Orikami Digital Health Products) app and a Fitbit Charge 2 wearable. In particular, we investigated if the distance measured by the s2MWT agreed with the distance measured using distance markers on the pavement. Furthermore, we investigated the first experiences of persons with MS with digital self-monitoring through smartphone apps and activity trackers by interviewing 7 study participants with MS as part of MS Self [16]. In this paper, we present the interview results for the s2MWT.

Methods

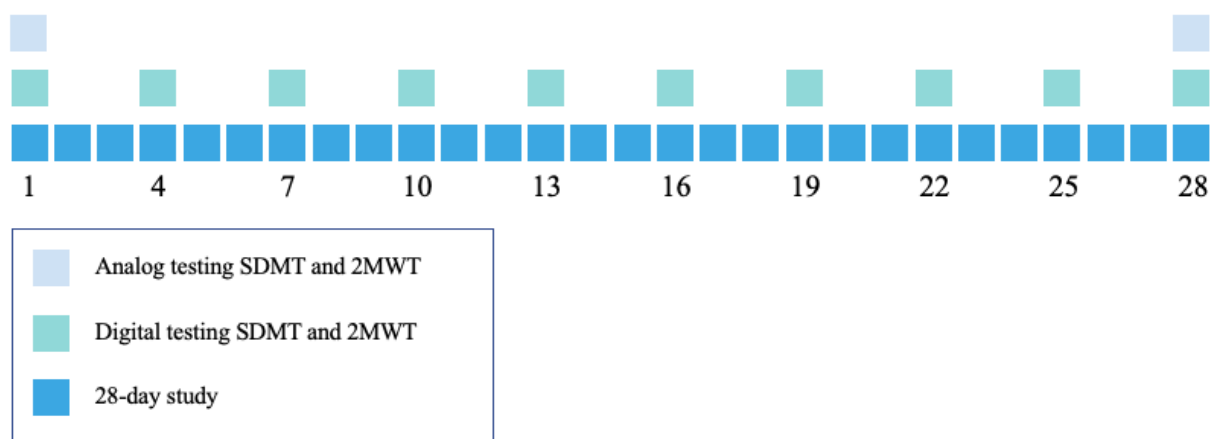
Study Design

We recently reported all relevant details about the study design, including the inclusion criteria, information about the recruitment of study participants, ethical approval and informed consent, and data collection in a publication about the validity and test-retest reliability of the smartphone variant of the Symbol Digit Modalities Test—as an assessment tool for cognitive processing speed—that is implemented in MS sherpa [17]. In summary, the study was performed on 25 persons with relapsing-remitting MS and 2 groups of healthy control (HC) participants (n=79). The HC participants in the first control group (HC matched, n=21) were matched to the persons with

MS with regard to age, gender, and education. The second control group (HC normative, n=58) was set up to determine the normal distribution for the smartphone test results. The app was installed on the study participants' smartphones during the first day of the study.

Figure 1 shows a schematic overview of the study design for the persons with MS and the participants in the HC-matched group. In total, 10 home assessments were planned for the persons with MS and the study participants in the HC-matched group in 28 days, with 3 days in between each test. Before the start of MS Self, 10 of the 25 persons with MS were randomly contacted by mail for the qualitative part of the study. Interviews were scheduled with 7 participants before and after the study, resulting in 14 interviews. More information about the interview methods is published elsewhere [16].

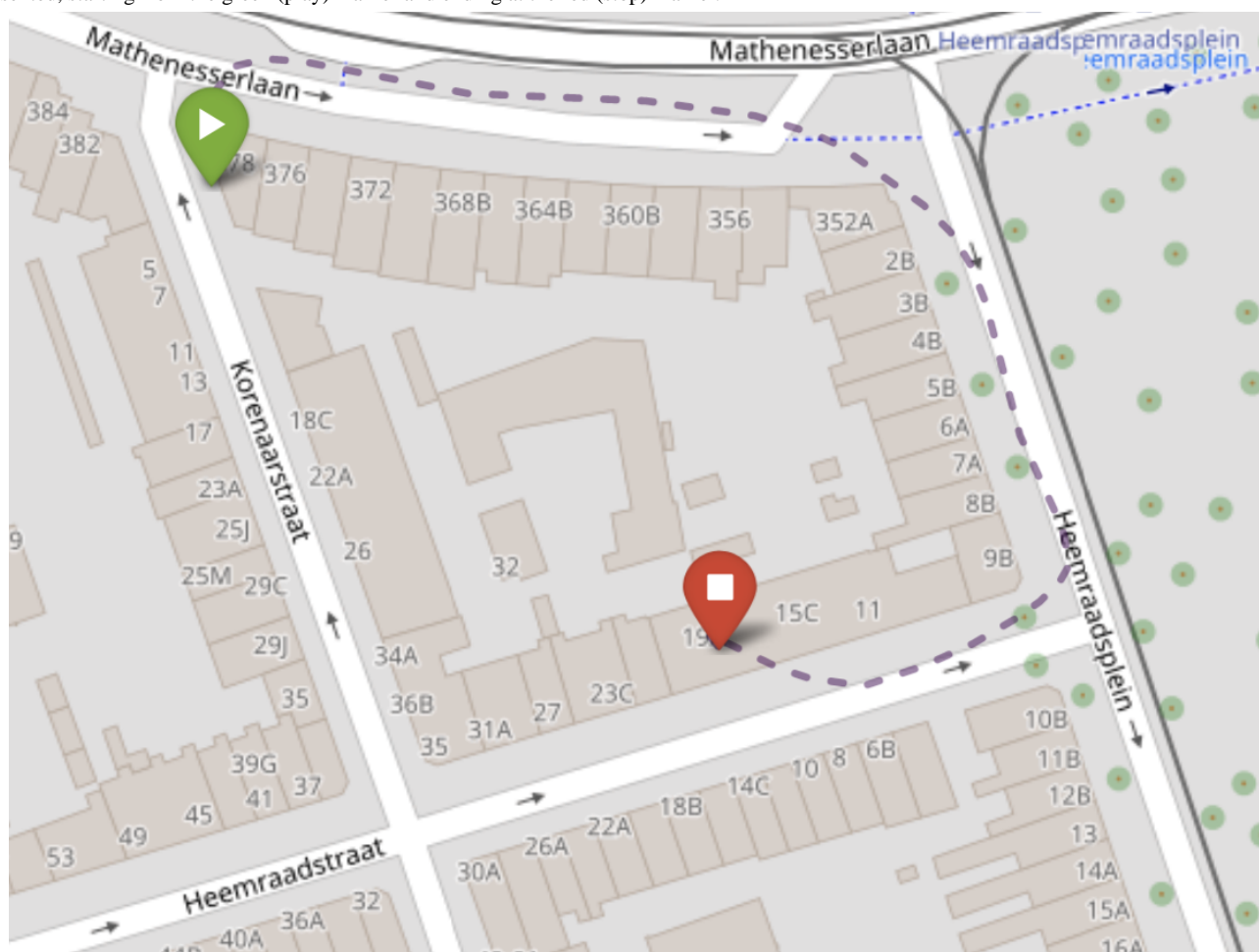
Figure 1. Overview of the study design and assessment scheme—the Symbol Digit Modalities Test (SDMT) and 2-Minute Walking Test (2MWT)—reproduced from van Oirschot et al [17].



The HC participants in the HC-normative group were instructed to perform the s2MWT 3 times in total, with 1 week in between the assessments. From these tests and from the 10 home assessments of the other study participants, the test-retest reliability of the s2MWT was determined. The concurrent validity of the s2MWT was determined from the comparison of the distance measured by the s2MWT with the distance measured using distance markers on the pavement during the s2MWT assessments—the 2MWTs that are described as *analog*

testing in Figure 1. Hereafter, these assessments are referred to as the validation assessments. Note that there are 2 validation assessments per study participant: one at the beginning and one at the end of the study. Figure 2 shows a map of the streets around the former premises of the Dutch National MS Foundation, Rotterdam, the Netherlands, on which the validation assessments took place, with an approximate path that was walked, reconstructed from one of the s2MWTs.

Figure 2. Street map of the block around the former premises of the Dutch National MS Foundation, Rotterdam, the Netherlands, on which the validation assessments took place. With a dashed line, the reconstructed path that was walked during one of the smartphone-based 2-Minute Walking Test is presented, starting from the green (play) marker and ending at the red (stop) marker.



Data Analysis

The distance walked during the s2MWT was calculated with a proprietary algorithm that applied a path reconstruction on the GPS data. This algorithm was tested on more than 10 different smartphones and optimized using various data sets held by the app manufacturer. Among the constraints applied on the reconstructed path is that it has to be possible to walk the path in 2 minutes.

In the data cleaning process, smartphone walking tests were removed when the data collected in a test were considered to be of low quality: (1) s2MWT duration <100 seconds; (2) s2MWT duration ≥ 140 seconds; (3) GPS data accuracy median ≥ 30 m; (4) GPS data accuracy SD ≥ 100 m; and (5) calculated distance walked <10 m.

For the test-retest analysis, we compared successive s2MWTs that were left after data cleaning and were less than 20 days apart.

Statistical Analysis

Unless mentioned otherwise, the statistical analysis was performed using SciPy (version 1.2.1) in combination with Python (version 3.6.7). Shapiro-Wilk tests were used to check if variables were normally distributed. If they were not, visual inspection led to the removal of at most 2 outliers, after which

a Shapiro-Wilk test confirmed a normal distribution. Therefore, the statistical tests mentioned could be applied. However, for the calculation of the Spearman rank correlations, we did not need to remove any outliers, as normality of the distributions is not a requirement for this calculation. *P* values <.05 were considered statistically significant for all statistical tests.

Two-sided *t* tests were conducted for the null hypothesis stating that the distance walked on the first s2MWT by persons with MS and distance walked by HC participants in both control groups have identical average (expected) values. A 2-sample Kolmogorov-Smirnov test was applied to the distributions of distance walked on the first s2MWT of the HC-matched group and HC-normative group to investigate if the 2 groups of HC participants had the same underlying distribution.

To investigate the concurrent validity of the s2MWT, we (1) conducted 2-sided *t* tests between the 2MWT distance and the s2MWT distance measured in the validation experiments; (2) calculated the Spearman rank correlation between the 2MWT distance and the s2MWT distance; (3) calculated the Spearman rank correlation between the EDSS score and the s2MWT distance; and (4) calculated the intraclass correlation coefficient (ICC) between the 2MWT distance and the s2MWT distance using a 2-way mixed effects model on absolute agreement for a single measurement: the ICC(A,1), following the nomenclature

of McGraw et al [21]. The ICCs were calculated using the R package *irr* version 0.84.1 in combination with R version 3.5.1. ICC values <0.20 were considered *poor*, values between 0.20 and 0.39 were considered *fair*, values between 0.40 and 0.59 were considered *moderate*, values between 0.60 and 0.79 were considered *good*, and values >0.80 were considered *very good* [22,23]. In all 3 analyses of the concurrent validity, we corrected for the fact that there were (at most) 2 validation assessments per user by replacing the individual observations by the subject mean [24].

Test-retest reliability was determined by calculating the ICC(A,1) between measurements at different times. We used the same acceptance criteria for the test-retest reliability ICCs as for the concurrent validity ICCs. These ICC values were used in combination with the pooled SD of the test and retest to determine the SEM and the smallest detectable change (SDC), using the formulas [25]:



Internal consistency was evaluated and quantified using Cronbach α , in which $\alpha > .7$ was defined to be acceptable [26,27]. The effect size, as measured by Cohen *d*, was determined to investigate the practice effect. Cohen *d* values <0.20 were considered small, values between 0.20 and 0.50 were considered medium, and values between 0.50 and 0.80 were considered large [28,29].

Results

Participant Demographics

The mean, median, and SD in the ages for the various groups are listed in Table 1, including the gender; number of participants in each education category (for the HC-normative group, no education information was collected); and the mean, median, and SD in the EDSS score (persons with MS). Even though having an EDSS score between 1.5 and 6.5 was one of the inclusion criteria, 1 person with MS with an EDSS score of 0 was enrolled in the study.

Table 1. Participant demographics.

Characteristics	Patients with relapsing-remitting multiple sclerosis (n=25)	HC ^a matched (n=21)	HC normative (n=58)
Age (years)			
Values, mean (SD)	40 (8)	37 (8)	34 (8)
Values, median	43	36	32
Gender			
Female, n	23	17	29
Male, n	2	4	29
Expanded Disability Status Scale			
Scores, mean (SD)	3.1 (1.4)	N/A ^b	N/A
Scores, median	3.0	N/A	N/A
Years since diagnosis			
Values, mean (SD)	6 (4.4)	N/A	N/A
Values, median	4	N/A	N/A

^aHC: healthy control.

^bN/A: not applicable.

Adherence and Number of Tests Done

In total, 104 study participants completed the assessments as scheduled, whereas 2 participants did not complete the study (1 person with MS and 1 HC participant in the HC-matched group). However, the validation assessments they did on the first day of the study could still be used to answer some of our research questions. The person with MS that dropped out of the study had an EDSS score of 3.5.

In total, 755 s2MWTs were done, which is 121 more than planned in the study protocol. This is because 91% (42/46) of the persons with MS and HC participants in the matched group continued to perform the tests after completing the 4 weeks of self-monitoring that was planned in the study protocol. This

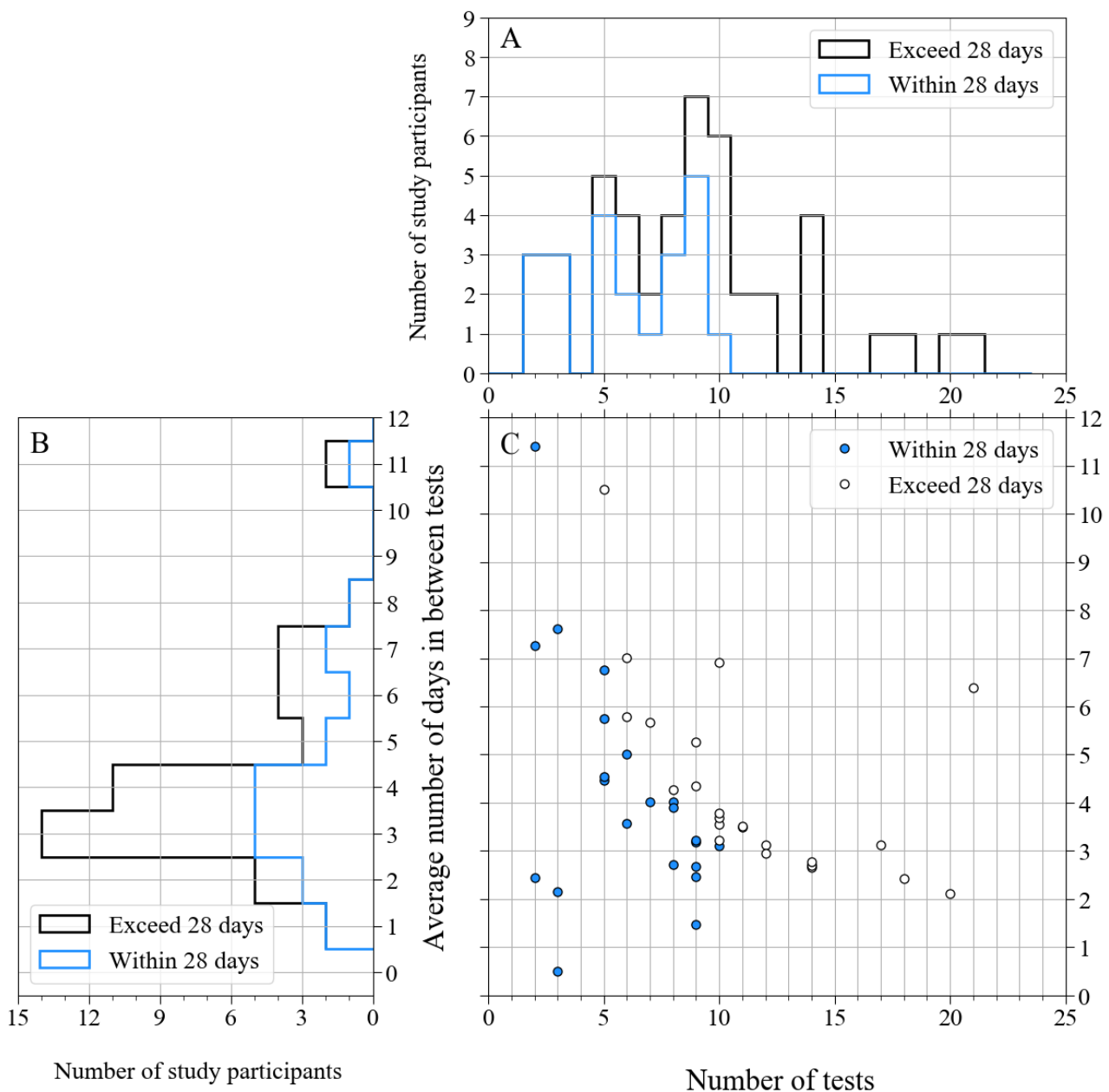
resulted in 72% (33/46) of these study participants completing at least 10 s2MWTs, the number that was planned in the study protocol, although mostly not with 3 days in between tests. The other 28% (13/46) of the persons with MS and HC participants in the matched group completed a total of 95 s2MWTs, which is, on average, more than 7 (SD 1.6) s2MWTs per person. In total, one could say that there was an s2MWT adherence rate of 92.4% (425/460) in these 2 groups. For the HCs in the normative group, who had a light study scheme with 3 s2MWTs in total, we only included participants who completed all 3 s2MWTs; therefore, we could not determine the adherence rate for this group.

One or more of the filtering criteria mentioned in the Data Analysis section was met in 23.3% (176/755) of all s2MWTs

that were completed. The data loss rates were 23.6% (87/369) for the persons with MS, 26.4% (56/212) for the participants in the HC-matched group, and 19% (33/174) for participants in the HC-normative group. In total, 15.8% (119/755) of the assessments were filtered out because the duration of the s2MWT was <100 or >140 seconds, 7.2% (54/755) of the assessments were filtered out because the GPS accuracy was too low, and 0.4% (3/755) of the assessments were filtered out because the predicted distance walked was ≤10 m.

Figure 3 shows the distribution of the number of s2MWTs performed by the persons with MS and HC participants that was left after data cleaning, the distribution of the average number of days between the tests per person with MS or HC participant in the matched group, and the relation between the number of tests and the average number of days in between tests in these groups.

Figure 3. (A) Histogram showing the distribution of the number of smartphone 2-Minute Walking Tests per study participant that remained after data cleaning and was used to determine the test-retest reliability. (B) Histogram showing the distribution of the average number of days in between these tests per study participant. (C) Scatter plot that shows the relation between the number of tests (horizontal axis) and the average number of days in between these tests (vertical axis) for all study participants. Filled circles correspond to participants whose tests were done within 28 days, and open circles correspond to those whose tests were done in more than 28 days. The blue lines in panels A and B correspond to these filled circles, and the black lines in panels A and B correspond to the open circles.



Even though there were more than 10 s2MWTs left after data cleaning that could be used to determine the test-retest reliability for 26% (12/46) of the persons with MS and HC participants in the matched group, we included at most the first 10

assessments for the test-retest reliability calculations, as planned in the study protocol. After data cleaning, 43.8% (160/365) of the consecutive s2MWTs were between 2 and 4 days apart, and 60.5% (221/365) were between 1 and 5 days apart. For 52%

(24/46) of the persons with MS and HC participants in the matched group, the timespan in which tests were completed exceeded the 4-week period that was planned in the study protocol.

There were missing smartphone data for 2 validation assessments on the first day of the study because of technical issues with the app, and 2 other validation assessments had missing smartphone data because of issues with the smartphone. Moreover, 1 study participant was not able to perform the validation assessment on the last day of the study because of the severity of MS; she had an EDSS score of 6.5 and walked <25 m on the first validation assessment.

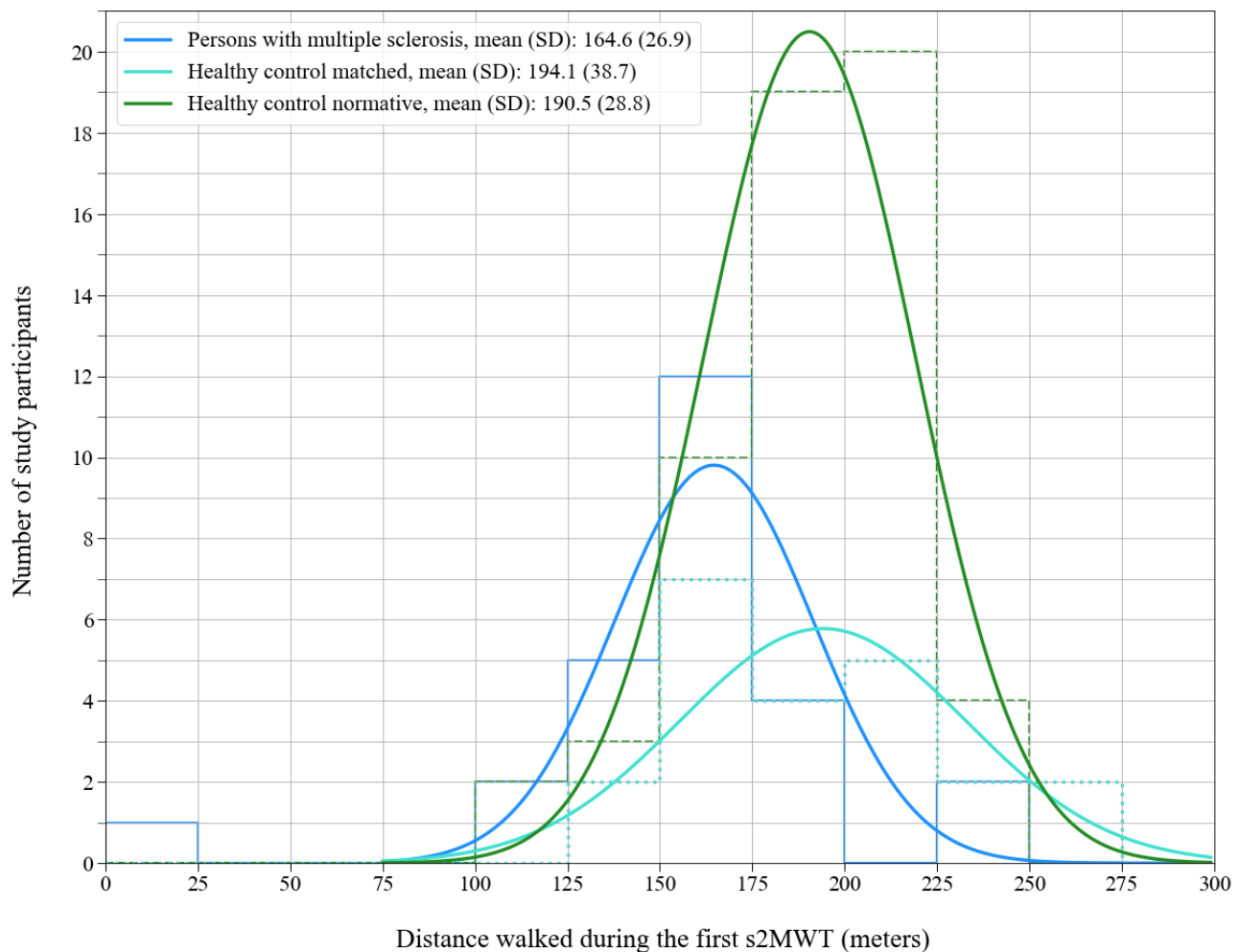
In the cleaning process, 35% (31/89) of the validation assessments were removed, including 1 of the 2 validation assessments of the dropouts because at least one of the filtering criteria mentioned in the Data Analysis section was met. We

averaged the distance walked in the 2 validation assessments of a study participant when both validation assessments of this study participant remained in our sample after data cleaning, as mentioned in the Statistical Analysis section. This resulted in 43 pairs of distance determinations (2MWT-s2MWT).

Distinction Between Persons With MS and HCs

Figure 4 shows the distance walked during the first valid s2MWT in the persons with MS and in the 2 control groups. Because one person with MS (with an EDSS score of 6.5) walked <25 m, the distribution of distances walked by the persons with MS was only normally distributed after removal of this data point ($P=.06$ on a Shapiro-Wilk test). The estimated distance walked was normally distributed for the HC-matched group and the HC-normative group because the Shapiro-Wilk tests on these distributions yielded $P=.48$ and $P=.24$, respectively.

Figure 4. Distributions of the distance walked on the first smartphone-based 2-Minute Walking Test (s2MWT) for the 3 groups in this study. The thin solid line shows the distribution for persons with multiple sclerosis, the dotted line shows the distribution for the healthy control-matched group, and the dashed line shows the distribution for the healthy control-normative group. The thick solid lines represent Gaussian fits to the distributions, of which the means (SDs) are shown in the legend.



The 2 groups of HC participants had the same underlying distribution as confirmed using a 2-sample Kolmogorov-Smirnov test (Kolmogorov-Smirnov statistic=0.21; $P=.44$). Independent 2-sample t tests between persons with MS versus the participants in the HC-matched

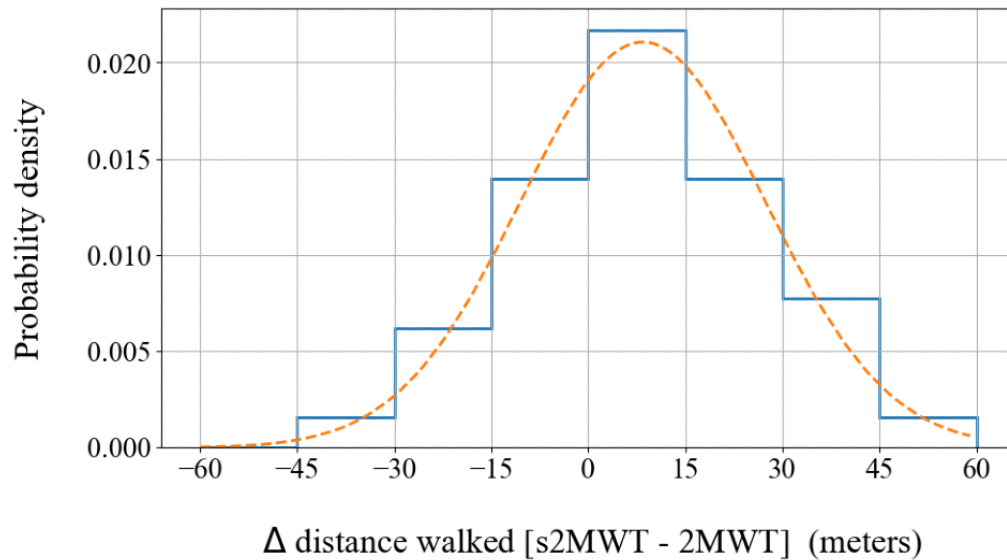
group and persons with MS versus the participants in the HC-normative group confirmed that the s2MWT can distinguish between persons with MS and HC participants at the group level ($P=.004$ and $P<.001$, respectively).

Concurrent Validity

A Shapiro-Wilk test on the distribution of the differences between the distance determined with the s2MWT and the distance measured using distance markers on the pavement

accepted normality ($P=.97$). This distribution is shown in Figure 5. The distance determined with the s2MWT was on average 8.43 m or 5% (SD 18.9 m or 11%) higher than the distance on the 2MWT ($n=43$). Here, the percentage is the difference divided by the 2MWT distance, times 100.

Figure 5. Distribution of differences between the measured distance walked by the app (smartphone-based 2-Minute Walking Test [s2MWT]) and the distance markers (2-Minute Walking Test [2MWT]). The dashed line represents a normal distribution.

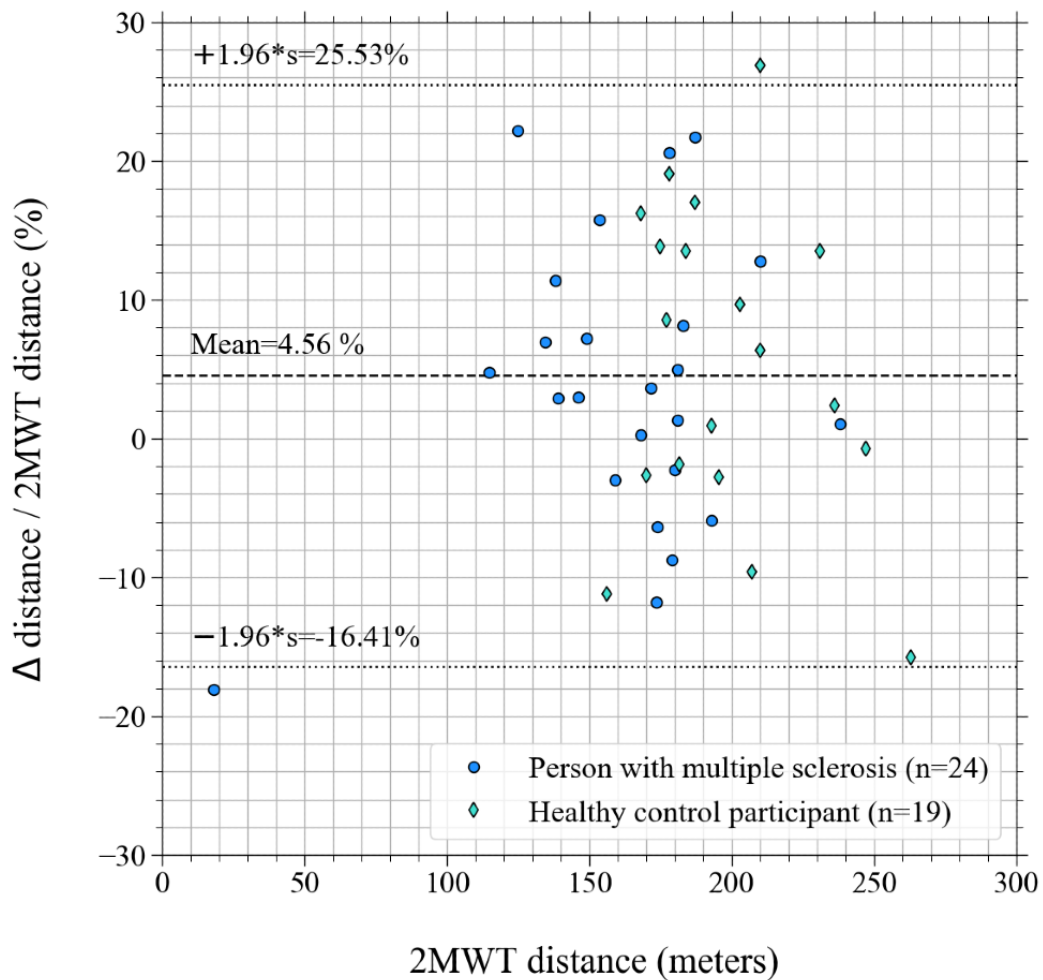


A dependent 2-sided t test between the 2MWT distance and the s2MWT distance measured in these validation assessments yielded a test statistic of -2.9 , with $P=.006$. Because this is below the significance level of 0.05, we must reject the null hypothesis of equal averages.

Figure 6 is a Bland-Altman plot of the 43 pairs of distance determinations derived from the validation assessments. The 2MWT distance is shown on the horizontal axis. On the vertical

axis of this plot, the percentage difference between the s2MWT and the 2MWT is shown, which is calculated as explained above. The mean percentage difference is shown as a horizontal dashed line, the 95% CI, at $1.96 \times SD$ around the mean difference is presented with 2 horizontal dotted lines. Distance determinations of persons with MS are presented with blue circles and those of HCs in the matched group, with green diamonds.

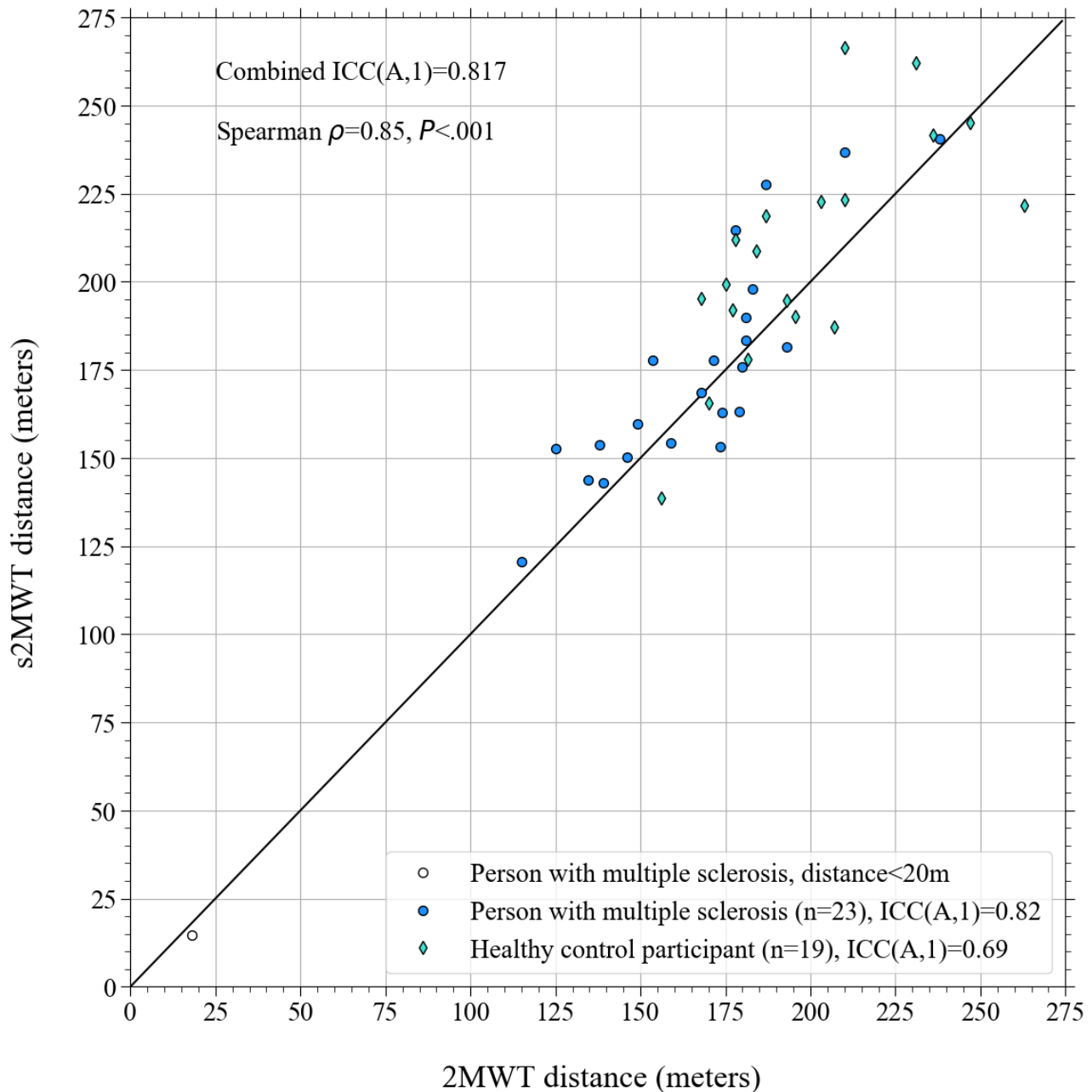
Figure 6. Bland Altman plot of differences between the measured "distance walked" by the app (smartphone-based 2-Minute Walking Test distance) and the distance markers (2-Minute Walking Test distance) expressed as percentages of the 2-Minute Walking Test distance (Δ distance / 2-Minute Walking Test distance) versus the 2-Minute Walking Test distance. The dashed line shows the mean percentage difference, the dotted lines show the 95% CI.



In Figure 7, the distance measured on the s2MWT is plotted against the distance measured on the 2MWT. ICC(A,1) values are shown for persons with MS and HCs separately (see legend in Figure 7) and for the combined data set, which is shown in

the top-left corner of Figure 7. The Spearman rank correlation and corresponding *P* value for the combined data set are also shown in the top-left corner of Figure 7.

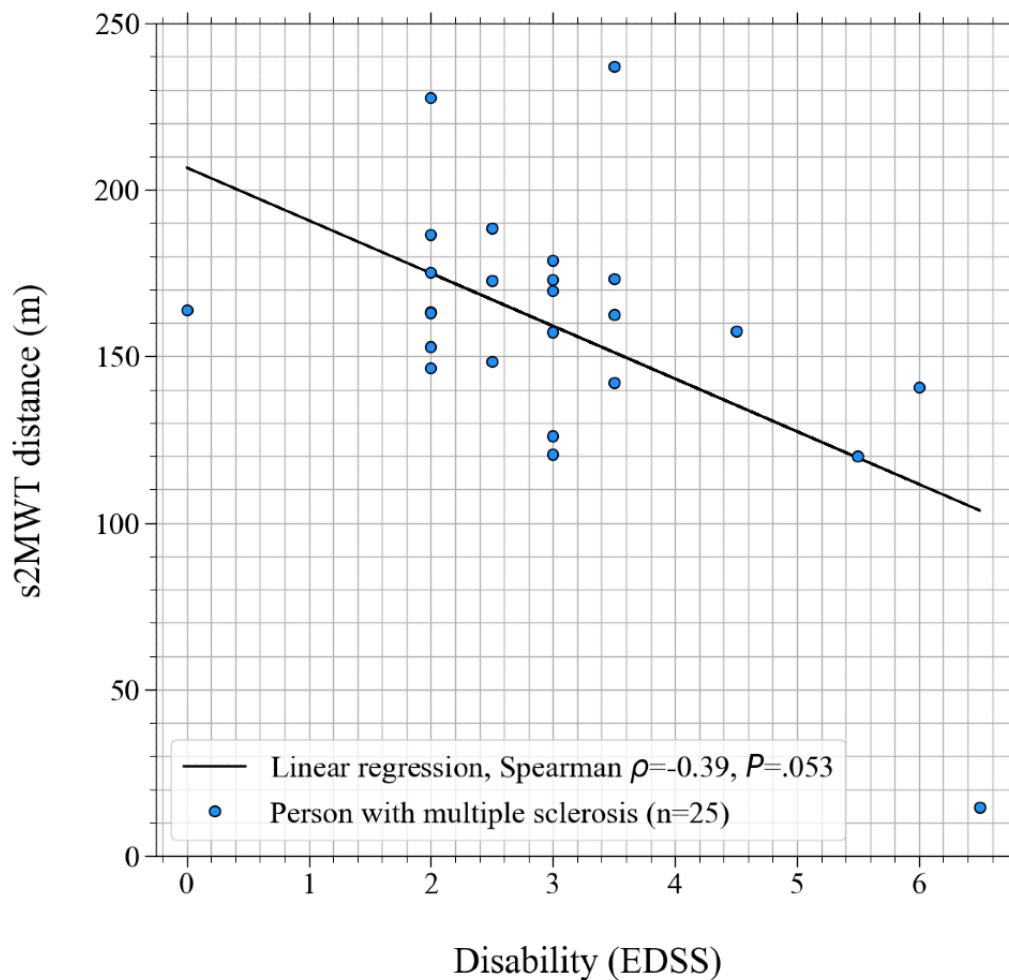
Figure 7. Scatter plot to show the ICC(A,1) values and the correlation (Spearman rho, upper left corner) between the measured 'distance walked' using distance markers (2-Minute Walking Test, horizontal axis) and the distance measured by the app (smartphone-based 2-Minute Walking Test, vertical axis). A 45 degree black solid line shows a 1:1 correlation.



The relation between the disability as measured by the EDSS and the s2MWT scores is presented in Figure 8. Distance determinations of persons with MS are presented with blue circles. A black solid linear regression line is overplotted. The

Spearman rank correlation and corresponding P values between the two variables are also shown in the bottom-left corner of Figure 8. We found a fair correlation that however failed to be statistically significant.

Figure 8. Relation between the distance walked on the first test done with the smartphone-based 2-Minute Walking Test and the Expanded Disability Status Scale score, for the 25 persons with multiple sclerosis that participated in this study.



Test-Retest Reliability

The ICCs(A,1) (with their 95% CI), Cronbach α , Cohen d , SEM, and SDC values that were derived from 9 test-retests of the s2MWT for the persons with MS and the participants in the HC-matched group are listed in Table 2, and the values of those derived from the 2 test-retests performed by the HC-normative group are shown in Table 3. The mean values of the ICC(A,1)

for persons with MS, participants in the HC-matched group, and participants in the HC-normative group were 0.649 (SD 0.150), 0.600 (SD 0.090), and 0.700 (SD 0.029), respectively. These indicate a good test-retest reliability. The corresponding average SDC values were 58.1 (SD 13.7) m, 61.8 (SD 15.0) m, and 51.1 (SD 0.1) m, which is 35% (58.1/165.5), 32% (61.8/193.5), and 27% (51.1/188.7) of the group's mean distance walked during test and retest, respectively.

Table 2. Test-retest reliability scores of the smartphone-based 2-Minute Walking Test for persons with multiple sclerosis and participants in the healthy control-matched group.

Test-retest	ICC ^a (A,1; 95% CI) of persons with MS ^b	ICC(A,1; 95% CI) of HC ^c participants	Cronbach α of persons with MS	Cronbach α of HC participants	Cohen <i>d</i> of persons with MS	Cohen <i>d</i> of HC participants	SEM of persons with MS (m)	SEM of HC participants (m)	SDC ^d of persons with MS (m)	SDC of HC participants (m)
1	0.462 (0.086 to 0.727)	0.554 (0.171 to 0.791)	.640	.708	0.301	0.150	24.0	30.6	66.4	84.7
2	0.509 (0.142 to 0.759)	0.589 (0.206 to 0.818)	.694	.742	0.369	0.219	24.0	28.6	66.6	79.3
3	0.624 (0.280 to 0.827)	0.723 (0.391 to 0.890)	.768	.863	0.185	0.145	23.2	17.6	64.2	48.8
4	0.657 (0.322 to 0.845)	0.439 (-0.057 to 0.756)	.787	.596	0.098	0.046	23.3	24.0	64.6	66.5
5	0.793 (0.547 to 0.913)	0.545 (-0.008 to 0.837)	.880	.690	0.035	0.066	15.9	23.0	44.2	63.7
6	0.873 (0.698 to 0.950)	0.659 (0.141 to 0.895)	.932	.788	0.114	0.198	12.3	23.2	34.1	64.4
7	0.823 (0.584 to 0.930)	0.633 (-0.031 to 0.905)	.912	.757	0.242	0.120	15.5	21.3	42.9	59.1
8	0.504 (0.059 to 0.781)	0.707 (-0.447 to 0.966)	.663	.794	0.154	0.010	25.5	12.8	70.6	35.4
9	0.595 (0.148 to 0.846)	0.553 (-0.773 to 0.986)	.761	.699	0.360	0.615	25.1	19.6	69.9	54.4
Mean	0.649 (0.318 to 0.842)	0.600 (-0.045 to 0.871)	.782	.734	0.206	0.174	21.0	22.3	58.1	61.8

^aICC: intraclass correlation coefficient.

^bMS: multiple sclerosis.

^cHC: healthy control.

^dSDC: smallest detectable change.

Table 3. Test-retest reliability of the smartphone-based 2-Minute Walking Test for the healthy control-normative group.

Test-retest	ICC ^a (A,1; 95% CI)	Cronbach α	Cohen <i>d</i>	SEM (m)	Smallest detectable change (m)
s2MWT test-retest 1	0.721 (0.548-0.835)	.835	0.031	18.4	51.0
s2MWT test-retest 2	0.680 (0.421-0.840)	.807	0.119	18.4	51.1
Mean	0.700 (0.480-0.837)	.821	0.075	18.4	51.1

^aICC: intraclass correlation coefficient.

We also calculated the ICC(A,1), Cronbach α , Cohen *d*, SEM, and SDC corresponding to the comparison of the average of the first 5 s2MWTs with the second 5 s2MWTs for persons with MS and the matched HCs. The results are shown in Table 4.

When averaged over 5 tests, the SDC values for persons with MS and the matched HCs were reduced to 16% (25.1/156.5) and 21% (40.7/193.5) of the group's mean distance walked, respectively.

Table 4. Test-retest reliability scores of the smartphone-based 2-Minute Walking Test based on the comparison of the mean value of the first 5 test scores with the mean value of the second 5 test scores for persons with multiple sclerosis and the healthy control participants in the matched group.

Group	ICC ^a (A,1; 95% CI)	Cronbach α	Cohen d	SEM (m)	Smallest detectable change (m)
Persons with multiple sclerosis	0.956 (0.898-0.981)	.977	0.028	9.0	25.1
Healthy control–matched	0.750 (0.381-0.915)	.859	0.220	14.7	40.7

^aICC: intraclass correlation coefficient.

The Cohen d values in Tables 2, 3, and 4 show that the practice effect is medium for persons with MS and small for HCs in the matched group when averaged over all test-retests. However, when comparing the average scores of the first 5 tests with those of the second 5 tests, the practice effect is small for persons with MS and medium for HCs in the matched group. For the HCs in the normative group, practice effects were small. When investigating the practice effect in the validation assessments, we found, on average, an increase of 10.6% (SD 12.5%) and 9.7% (SD 7.8%) in 2MWT distance walked for persons with MS and distance walked by HC participants, respectively. With corresponding Cohen d values of 0.66 (persons with MS) and 0.59 (HC-matched participants), this is considered a large practice effect.

Interview Results

Out of 7 participants with MS, 5 participants with MS who were interviewed about their experiences with the smartphone app and Fitbit activity tracker in general experienced some technical difficulties with the test, such as issues with the GPS signal or that the test suddenly stopped. Respondents were frustrated about these issues. Furthermore, 3 respondents mentioned that they had more or less stable s2MWT results during the study, whereas for 4 respondents, the results showed more fluctuation. One respondent explicitly stated that being faced with these fluctuations was emotionally confronting to her. Moreover, 3 respondents experienced a competitive element in the s2MWT—feeling the pressure to reach the same distance every time. Of these, 1 respondent wanted to *improve* her score by walking a little faster. Difficulties in making the s2MWT a routine element of their daily life were reported by 3 respondents. For instance, 1 respondent expressed annoyance about having to perform the test at specific moments of the week, whereas another respondent sometimes forgot to do the test, which made her feel guilty. However, in general the interview respondents expressed that the instructions for the s2MWT were clear and that the test was easy to perform.

Discussion

Principal Findings

We found that the s2MWT can distinguish between persons with MS and HC participants at the group level. The estimated distance walked on the s2MWT is, on average, 5% (SD 11%) higher than the distance measured using distance markers. Therefore, we cannot conclude that the average s2MWT distance is equal to the average 2MWT distance, which is also reflected in the outcome of a dependent 2-sided t test between the 2 distributions ($P < .05$). The duration of the s2MWTs of the

persons with MS and matched HCs was, on average, 2.86 (SD 5.74) seconds too long, which could explain why the s2MWT overestimates the distance walked. However, the distances measured with the s2MWT highly correlated with the 2MWT distances; we found a Spearman rank correlation of 0.85. Furthermore, a very good concurrent validity for persons with MS and a good concurrent validity for HC participants was established, given the ICC(A,1) values of 0.82 and 0.69 between the s2MWT distance and the 2MWT distance, respectively.

The correlation between the s2MWT score and disability as measured by the EDSS was fair but not significant. However, per the definition of the EDSS, persons with MS having EDSS scores of 4 or lower have full ambulation (including the ability to walk without aid or rest for some 500 m) [3]. As there were only 4 study participants with EDSS scores >4 , it is not strange that the correlation that we found was not significant. Furthermore, if we had not included the person with the EDSS score of 0, we would have found a significant ($P = .049$) Spearman rank correlation of -0.41 between EDSS and the distance walked on the first s2MWT.

The 9 test-retest analyses showed that the test-retest reliability was good for the persons with MS (average ICC[A,1] 0.649, SD 0.150) and matched HCs (average ICC[A,1] 0.600, SD 0.090). In addition, the test-retest reliability we found in the HC-normative group was good, with an average ICC(A,1) value of 2 test-retest analyses of 0.700 (SD 0.029). Practice effects between consecutive tests were, on average, medium for persons with MS and small for HCs in the matched group. However, the practice effect between the average score of the first 5 s2MWTs compared with the average score of the second 5 s2MWTs was small for persons with MS and medium for HCs in the matched group. The practice effect between the 2 validation assessments was large. The study participants may have felt pressure to reach the same distance as during the first validation assessment, which could have resulted in faster walking speeds during the second validation assessment. On average, there were 65 (SD 44) days in between the 2 validation assessments.

We derived an SDC of 58.1 m and 61.8 m for persons with MS and matched HCs, respectively, from the individual s2MWT test-retest assessments. These values seem rather large given the respective typical walked distances of 165 m and 194 m of these groups. Large day-to-day variations in the distance walked during the s2MWT and the s2MWT measurement error contributed to the large SDC. However, a smaller SDC can be obtained by averaging over multiple measurements. Indeed, from comparing the average distance walked in the first 5 s2MWTs with the average distance walked in the second 5

s2MWTs, we derived SDC values of 25.1 m and 40.7 m for persons with MS and matched HCs, respectively. This finding implies that changes in walking speed that result in a more than 25.1 m change on the s2MWT score, when averaged over 5 measurements, are above day-to-day variations and measurement noise.

It is important to distinguish the SDC from the minimally important clinical difference (MICD). The MICD is the smallest change in test score that is perceived as important by patients and clinicians, whereas the SDC is the smallest change that can be detected beyond measurement error. Preferably, a measurement instrument has an SDC that is smaller than the MICD, so that all clinically relevant changes can be distinguished from measurement error. For the 2MWT, a MICD of <10 m was found after MS rehabilitation [30], but a change of 20% is generally considered an MICD for the T25FW [31]—another walking test that is often used in MS—that highly correlates with the 2MWT [6]. Although the SDC that we find is larger than 10 m, the 25.1 m that we found after averaging 5 measurements correspond to 16% of the group's mean distance walked, which is less than 20%.

The interview respondents shared some important considerations. Aside from technical issues that need to be solved, performing the test can evoke emotional responses. For instance, patients can feel confronted by the test results, experience pressure to reach a certain distance, or feel annoyed by having to perform the test regularly. However, the interview respondents also expressed that the instructions for the s2MWT were clear and that the test was easy to perform. Furthermore, in clinical practice, the s2MWT is expected to be scheduled less often than every 3 days, which was the test frequency applied in this study. If the s2MWTs are scheduled, for example, weekly, always at the same day of the week and at the same time of the day chosen by the person with MS, it should be easier to make self-monitoring walking speed a routine element of daily life.

Limitations and Future Work

The 2MWT is a relevant outcome measure not only for persons with MS but in a variety of health conditions such as chronic obstructive pulmonary disease [32], lower limb amputation [33], cardiovascular disease [34], osteoarthritis [35], Parkinson disease [36], and Alzheimer disease [37]. This study limits itself to persons with MS, and the validity of the s2MWT for applications outside MS will still have to be demonstrated. Furthermore, 84% (21/25) of the persons with MS participating in this study had an EDSS score below 4. As can be seen in Figures 5 and 6, all persons with MS walked ≥ 100 m in the validation assessments, apart from 1 study participant (who had an EDSS score of 6.5; Figure 8). Therefore, this validation of the s2MWT limits itself to persons with MS who are able to walk ≥ 100 m in 2 minutes. The validity of the s2MWT for slow walkers, that is, those who walk ≤ 100 m in 2 minutes, has to be demonstrated in a different validation study.

For this analysis, we had to limit ourselves to 76.7% (579/755) of the s2MWT assessments done during MS Self. Most assessments that were filtered out did not pass the acceptance criteria that the s2MWT duration should be 120 seconds within

a 20-second margin. The assessment duration is derived from the time difference between the first and the last GPS data point. The GPS data collection frequency is typically 1 data point per second or even per 2 seconds; therefore, the s2MWT assessment duration is often 1 or sometimes even 2 seconds ≤ 2 minutes. s2MWT durations that have much larger deviations from 120 seconds than 2 seconds must have been the result of data collection deficiency. This data collection deficiency may have arisen from memory loss or a clock synchronization error when the app was unintentionally running in the background. The 20-second margin that we considered acceptable here is far from ideal. Assessments that are on the border of being acceptable (eg, an assessment with a duration of 105 seconds or an assessment with a duration of 139 seconds) are expected to result in a poor estimation of the distance walked in 2 minutes. The data collection during s2MWT assessments was improved after this study, and the improved s2MWT is currently implemented in the MS sherpa app. It is expected that this new s2MWT has a higher concurrent validity and test-retest reliability than the s2MWT that was used in this study. Preliminary results from the APPS MS (Assessing fatigue, disease activity and Progression through smartPhone Surveillance in Multiple Sclerosis) study confirm that the test-retest reliability has indeed improved [18].

One should keep in mind that the s2MWT was performed on the study participant's phone, and technical specifications, in particular GPS accuracy, may not be equivalent for all devices. Although this study was not set up to compare s2MWT performances of different phone types, we have not found large differences between phone types in preclinical studies. Here, we have partially accounted for the variability in phone types in the data cleaning process, where we filtered out assessments with low GPS accuracy.

Another limitation of this study and an opportunity for future improvements is the algorithm that was used to calculate the distance walked. The version used for this paper makes use of the GPS data alone, while accelerometer data was also collected during the s2MWTs in this study. GPS gives information about where a person has walked, while the accelerometer provides information about the forces that are exerted on the mobile phone during the 2MWT, including gravity. The accelerometer comes from an inertial measurement unit that is embedded in the phone. Future work can improve the algorithm by making use of the accelerometer data and data from various other sensors of the inertial measurement unit, such as angular velocity from a gyroscope. These additional sources of data allow for various other kinds of information to be extracted, such as balance and gait characteristics [8-13]. Furthermore, inertial measurement unit data have been shown to be suitable for walking distance estimation, for example, using machine learning or dead reckoning [38,39]. The reason to prefer an alternative to a purely GPS-based solution is that GPS has inherent inaccuracies because of the dependency of satellites in orbit. Anything such as large buildings or overpasses can significantly reduce the accuracy of the GPS, resulting in inaccurate distance estimations. The strengths of both data sources combined can positively influence the reliability of the measurements.

In the s2MWT instructions during MS Self, it was not requested to walk the same path every time the test is done. If the test and retest were not done on the same location, the test result was likely to be affected because of environmental effects, for example, in an uphill walk, the user was expected to reach a shorter distance than in a flat walk. In additional analyses, we observed that walking a straight line was also preferable to reach more accurate distance calculations. Therefore, we improved the instructions for the users in the s2MWT that is currently implemented in the MS sherpa app—they should walk in a straight line as much as possible and try to walk the same route every time they do the test.

In an MS clinic, a 2MWT assessment outcome is typically established after averaging the distances walked in two or three 2MWTs. In this study, we scheduled one s2MWT at a time for the convenience of the persons with MS performing the assessments. Averaging over repeated smartphone assessments may also improve the accuracy of the test. This should also be kept in mind when comparing the test-retest reliability of various walking tests in the literature.

The 2MWT distance estimations in the validation assessments of MS Self may be up to a few meters off because of small deviations in the walked paths with respect to the path set out by the distance markers. In future studies that include validation experiments such as those in MS Self, we propose to use something akin to a measuring wheel, also known as a surveyor's wheel, to determine the actual distance walked during a 2MWT by walking alongside the person being tested. We expect that the distance walked in real life settings can be determined very accurately with a measuring wheel. Furthermore, it would be easy to choose a different path for various validation assessments instead of following the same distance markers as closely as possible each time.

As mentioned in the Principal Findings section, there were, on average, 65 (SD 44) days in between the 2 validation assessments, and as shown in Figure 3, some participants had more than 130 days between the 2 study visits. Although this does not affect the smartphone test-retest reliability, as we only compared successive s2MWTs that were less than 20 days apart for each user, it should be kept in mind when interpreting the practice effect in the validation assessments. Furthermore, we did not investigate whether the persons with MS experienced any relapse during the follow-up nor did we investigate whether they took any new medication that could have influenced their walking performances, which would affect the test-retest reliability. However, we did register that the occurrence of a relapse was the reason to drop out of the study for one of the included persons with MS.

In a future study, we will investigate how frequently the s2MWT should be performed to obtain a clear picture of the walking speed of a person with MS over time. Furthermore, we will investigate how monitoring walking speed in a patients' home environment could help in making clinical decisions, for example, in the evaluation of the effectiveness of a medicine

used to improve walking ability in persons with MS, such as fampridine, or for measuring disease activity or disease progression in MS. One can imagine that this digital biomarker could help to early predict a transition to the secondary progressive phase or to detect suboptimal treatment response. Furthermore, this digital biomarker could potentially be used for evaluating of the effect of MS rehabilitation.

Comparison With Previous Work

Although much of the literature is available on the 2MWT as a relevant tool to measure walking speed in persons with MS [6,40-44], little has been published about 2MWTs that can be self-administered on a smartphone. The Floodlight smartphone monitoring app for persons with MS, which was developed by Roche, contains an s2MWT, that was shown to moderately correlate with T25FW time [45]. This has recently also been shown for the s2MWT that is implemented in MS sherpa [18]. The MSCopilot smartphone monitoring app for persons with MS contains another alternative to self-assessed walking speed. However, results on the performance of this smartphone walking test have not been individually reported but only for the combined digital Multiple Sclerosis Functional Composite assessment scores [15].

It is already known that self-monitoring can evoke strong emotions and sentiments [46]. It is crucial to be aware of the potential burden of self-monitoring to patients, as a significant user burden leads to unwillingness to use these technologies [47]. Therefore, it is important to explore how digital self-monitoring tools could be developed in such a way that the burden is reduced [46,47].

Conclusions

This study shows a good concurrent validity of the s2MWT because the ICCs(A,1) between the 2MWT and the s2MWT for persons with MS and HC participants were 0.82 and 0.69, respectively. The distance determined with the s2MWT is, on average, 4.56% (SD 10.7%) larger than the distance measured using distance markers on the pavement. It is expected that this can largely be attributed to the s2MWT assessment durations that were, on average, 2.86 (SD 5.74) seconds too long—an artifact of the s2MWT that was used in this study that is no longer present in the s2MWT implemented in MS sherpa. The s2MWT has a good test-retest reliability because the ICC(A,1) values averaged over all test-retests performed by the persons with MS and both groups of HCs were in the range 0.6-0.7. We conclude that the s2MWT can be used to measure the distance that persons with MS walk in 2 minutes in the outdoors near their home. Clinical studies and clinical practice can benefit from this, as it allows the collection of real-world evidence for interventions aimed to improve the walking speed of a person with MS. Furthermore, a frequent assessment of walking speed in the home environment of a person with MS may improve clinical decision-making because it provides health care providers with quantitative information about changes in their patients' walking speed.

Acknowledgments

The study received financial support from the European Regional Development Fund (EFRO OP Oost Project 00217) and the Dutch National Multiple Sclerosis Foundation, Rotterdam, the Netherlands.

Conflicts of Interest

PO, FD, and RE are employed by Orikami Digital Health Products. BT is the founder and owner of Orikami Digital Health Products. PJJ is adviser to Orikami Digital Health Products and has received honoraria from Bayer Netherlands for consultancy activities.

References

1. Embrey N. Multiple sclerosis: managing a complex neurological disease. *Nurs Stand* 2014 Nov 18;29(11):49-58 [FREE Full text] [doi: [10.7748/ns.29.11.49.e9190](https://doi.org/10.7748/ns.29.11.49.e9190)] [Medline: [25388738](https://pubmed.ncbi.nlm.nih.gov/25388738/)]
2. Bethoux F, Bennett S. Evaluating walking in patients with multiple sclerosis: which assessment tools are useful in clinical practice? *Int J MS Care* 2011;13(1):4-14 [FREE Full text] [doi: [10.7224/1537-2073-13.1.4](https://doi.org/10.7224/1537-2073-13.1.4)] [Medline: [24453700](https://pubmed.ncbi.nlm.nih.gov/24453700/)]
3. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983 Nov;33(11):1444-1452 [FREE Full text] [doi: [10.1212/wnl.33.11.1444](https://doi.org/10.1212/wnl.33.11.1444)] [Medline: [6685237](https://pubmed.ncbi.nlm.nih.gov/6685237/)]
4. Kahraman T, Savci S, Coskuner-Poyraz E, Ozakbas S, Idiman E. Determinants of physical activity in minimally impaired people with multiple sclerosis. *Clin Neurol Neurosurg* 2015 Nov;138:20-24 [FREE Full text] [doi: [10.1016/j.clineuro.2015.07.018](https://doi.org/10.1016/j.clineuro.2015.07.018)] [Medline: [26264722](https://pubmed.ncbi.nlm.nih.gov/26264722/)]
5. Pokryszko-Dragan A, Marschollek K, Chojko A, Karasek M, Kardyś A, Marschollek P, et al. Social participation of patients with multiple sclerosis. *Adv Clin Exp Med* 2020 Apr;29(4):469-473 [FREE Full text] [doi: [10.17219/acem/115237](https://doi.org/10.17219/acem/115237)] [Medline: [32343889](https://pubmed.ncbi.nlm.nih.gov/32343889/)]
6. Bennett S, Bromley L, Fisher N, Tomita M, Niewczyk P. Validity and reliability of four clinical gait measures in patients with multiple sclerosis. *Int J MS Care* 2017;19(5):247-252 [FREE Full text] [doi: [10.7224/1537-2073.2015-006](https://doi.org/10.7224/1537-2073.2015-006)] [Medline: [29070965](https://pubmed.ncbi.nlm.nih.gov/29070965/)]
7. Shamma L, Zentek T, von Haaren B, Schlesinger S, Hey S, Rashid A. Home-based system for physical activity monitoring in patients with multiple sclerosis (Pilot study). *BioMed Eng OnLine* 2014 Feb 6;13(1):10 [FREE Full text] [doi: [10.1186/1475-925x-13-10](https://doi.org/10.1186/1475-925x-13-10)]
8. Supratak A, Datta G, Gafson AR, Gafson AR, Guo Y, Matthews PM. Remote monitoring in the home validates clinical gait measures for multiple sclerosis. *Front Neurol* 2018;9:561 [FREE Full text] [doi: [10.3389/fneur.2018.00561](https://doi.org/10.3389/fneur.2018.00561)] [Medline: [30057565](https://pubmed.ncbi.nlm.nih.gov/30057565/)]
9. Midaglia L, Mulero P, Montalban X, Graves J, Hauser S, Julian L, et al. Adherence and satisfaction of smartphone- and smartwatch-based remote active testing and passive monitoring in people with multiple sclerosis: nonrandomized interventional feasibility study. *J Med Internet Res* 2019 Aug 30;21(8):e14863 [FREE Full text] [doi: [10.2196/14863](https://doi.org/10.2196/14863)]
10. Creagh AP, Simillion C, Bourke AK, Scotland A, Lipsmeier F, Bernasconi C, et al. Smartphone- and smartwatch-based remote characterisation of ambulation in multiple sclerosis during the two-minute walk test. *IEEE J Biomed Health Inform* 2021 Mar;25(3):838-849 [FREE Full text] [doi: [10.1109/JBHI.2020.2998187](https://doi.org/10.1109/JBHI.2020.2998187)] [Medline: [32750915](https://pubmed.ncbi.nlm.nih.gov/32750915/)]
11. Bourke A, Scotland A, Lipsmeier F, Gossens C, Lindemann M. Gait characteristics harvested during a smartphone-based self-administered 2-minute walk test in people with multiple sclerosis: test-retest reliability and minimum detectable change. *Sensors (Basel)* 2020 Oct 19;20(20):5906 [FREE Full text] [doi: [10.3390/s20205906](https://doi.org/10.3390/s20205906)] [Medline: [33086734](https://pubmed.ncbi.nlm.nih.gov/33086734/)]
12. Abou L, Wong E, Peters J, Dossou MS, Sosnoff JJ, Rice LA. Smartphone applications to assess gait and postural control in people with multiple sclerosis: a systematic review. *Mult Scler Relat Disord* 2021 Jun;51:102943. [doi: [10.1016/j.msard.2021.102943](https://doi.org/10.1016/j.msard.2021.102943)] [Medline: [33873026](https://pubmed.ncbi.nlm.nih.gov/33873026/)]
13. Hsieh KL, Sosnoff JJ. Smartphone accelerometry to assess postural control in individuals with multiple sclerosis. *Gait Posture* 2021 Feb;84:114-119. [doi: [10.1016/j.gaitpost.2020.11.011](https://doi.org/10.1016/j.gaitpost.2020.11.011)] [Medline: [33307327](https://pubmed.ncbi.nlm.nih.gov/33307327/)]
14. Pratap A, Grant D, Vegesna A, Tummalacherla M, Cohan S, Deshpande C, et al. Evaluating the utility of smartphone-based sensor assessments in persons with multiple sclerosis in the real-world using an app (elevateMS): observational, prospective pilot digital health study. *JMIR Mhealth Uhealth* 2020 Oct 27;8(10):e22108 [FREE Full text] [doi: [10.2196/22108](https://doi.org/10.2196/22108)] [Medline: [33107827](https://pubmed.ncbi.nlm.nih.gov/33107827/)]
15. Maillart E, Labauge P, Cohen M, Maarouf A, Vukusic S, Donz c C, et al. MSCopilot, a new multiple sclerosis self-assessment digital solution: results of a comparative study versus standard tests. *Eur J Neurol* 2020 Mar;27(3):429-436 [FREE Full text] [doi: [10.1111/ene.14091](https://doi.org/10.1111/ene.14091)] [Medline: [31538396](https://pubmed.ncbi.nlm.nih.gov/31538396/)]
16. Wendrich K, van Oirschot P, Martens M, Heerings M, Jongen P, Krabbenborg L. Toward digital self-monitoring of multiple sclerosis: investigating first experiences, needs, and wishes of people with MS. *Int J MS Care* 2019 Nov;21(6):282-291 [FREE Full text] [doi: [10.7224/1537-2073.2018-083](https://doi.org/10.7224/1537-2073.2018-083)] [Medline: [31889935](https://pubmed.ncbi.nlm.nih.gov/31889935/)]
17. van Oirschot P, Heerings M, Wendrich K, den Teuling B, Martens M, Jongen P. Symbol digit modalities test variant in a smartphone app for persons with multiple sclerosis: validation study. *JMIR Mhealth Uhealth* 2020 Oct 05;8(10):e18160 [FREE Full text] [doi: [10.2196/18160](https://doi.org/10.2196/18160)] [Medline: [33016886](https://pubmed.ncbi.nlm.nih.gov/33016886/)]

18. Lam K, van Oirschot P, Dorssers F, den Teuling BD, Uitdehaag B, De Groot V, et al. Reliability, concurrent and ecological validity of smartphone-based cognition and walking tests. In: Poster Presentation at MS Virtual 2020, the 8th Joint ACTRIMS-ECTRIMS Meeting. 2020 Presented at: MS Virtual 2020, the 8th Joint ACTRIMS-ECTRIMS Meeting; September 11-13, 2020; Online URL: <https://tinyurl.com/fnwufmmc>
19. Lam K, van Oirschot P, den Teuling B, Hulst HE, de Jong BA, Uitdehaag B, et al. Reliability, construct and concurrent validity of a smartphone-based cognition test in multiple sclerosis. *Mult Scler* 2021 May 26;13524585211018103 [FREE Full text] [doi: [10.1177/13524585211018103](https://doi.org/10.1177/13524585211018103)] [Medline: [34037472](https://pubmed.ncbi.nlm.nih.gov/34037472/)]
20. MS sherpa. URL: <https://www.mssherpa.com/> [accessed 2021-10-08]
21. McGraw K, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1(1):30-46 [FREE Full text] [doi: [10.1037/1082-989x.1.1.30](https://doi.org/10.1037/1082-989x.1.1.30)]
22. Bushnell CD, Johnston DC, Goldstein LB. Retrospective assessment of initial stroke severity: comparison of the NIH Stroke Scale and the Canadian Neurological Scale. *Stroke* 2001 Mar;32(3):656-660 [FREE Full text] [doi: [10.1161/01.str.32.3.656](https://doi.org/10.1161/01.str.32.3.656)] [Medline: [11239183](https://pubmed.ncbi.nlm.nih.gov/11239183/)]
23. Chien CW, Hu MH, Tang PF, Sheu CF, Hsieh CL. A comparison of psychometric properties of the smart balance master system and the postural assessment scale for stroke in people who have had mild stroke. *Arch Phys Med Rehabil* 2007 Mar;88(3):374-380 [FREE Full text] [doi: [10.1016/j.apmr.2006.11.019](https://doi.org/10.1016/j.apmr.2006.11.019)] [Medline: [17321832](https://pubmed.ncbi.nlm.nih.gov/17321832/)]
24. Bland J, Altman D. Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. *Br Med J* 1995 Mar 11;310(6980):633 [FREE Full text] [doi: [10.1136/bmj.310.6980.633](https://doi.org/10.1136/bmj.310.6980.633)] [Medline: [7703752](https://pubmed.ncbi.nlm.nih.gov/7703752/)]
25. de Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine - A Practical Guide*. New York, United States: Cambridge University Press; 2011.
26. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951 Sep;16(3):297-334 [FREE Full text] [doi: [10.1007/bf02310555](https://doi.org/10.1007/bf02310555)]
27. Nunnally JC. *Psychometric Theory*. New York: McGraw-Hill; 1978:1-701.
28. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988:1-400.
29. Sawilowsky SS. New effect size rules of thumb. *J Mod App Stat Meth* 2009 Nov 01;8(2):597-599. [doi: [10.22237/jmasm/1257035100](https://doi.org/10.22237/jmasm/1257035100)]
30. Baert I, Freeman J, Smedal T, Dalgas U, Romberg A, Kalron A, et al. Responsiveness and clinically meaningful improvement, according to disability level, of five walking measures after rehabilitation in multiple sclerosis: a European multicenter study. *Neurorehabil Neural Repair* 2014 Sep 06;28(7):621-631. [doi: [10.1177/1545968314521010](https://doi.org/10.1177/1545968314521010)] [Medline: [24503204](https://pubmed.ncbi.nlm.nih.gov/24503204/)]
31. Hoogervorst EL, Kalkers NF, Cutter GR, Uitdehaag BM, Polman CH. The patient's perception of a (reliable) change in the Multiple Sclerosis Functional Composite. *Mult Scler* 2004 Feb;10(1):55-60. [doi: [10.1191/1352458504ms972oa](https://doi.org/10.1191/1352458504ms972oa)] [Medline: [14760953](https://pubmed.ncbi.nlm.nih.gov/14760953/)]
32. Johnston K, Potter A, Phillips A. Minimal important difference and responsiveness of 2-minute walk test performance in people with COPD undergoing pulmonary rehabilitation. *COPD* 2017 Oct 9;Volume 12:2849-2857 [FREE Full text] [doi: [10.2147/copd.s143179](https://doi.org/10.2147/copd.s143179)]
33. Brooks D, Parsons J, Hunter J, Devlin M, Walker J. The 2-minute walk test as a measure of functional improvement in persons with lower limb amputation. *Arch Phys Med Rehabil* 2001 Oct;82(10):1478-1483 [FREE Full text] [doi: [10.1053/apmr.2001.25153](https://doi.org/10.1053/apmr.2001.25153)] [Medline: [11588757](https://pubmed.ncbi.nlm.nih.gov/11588757/)]
34. Brooks D, Parsons J, Tran D, Jeng B, Gorczyca B, Newton J, et al. The two-minute walk test as a measure of functional capacity in cardiac surgery patients. *Arch Phys Med Rehabil* 2004 Sep;85(9):1525-1530 [FREE Full text] [doi: [10.1016/j.apmr.2004.01.023](https://doi.org/10.1016/j.apmr.2004.01.023)] [Medline: [15375829](https://pubmed.ncbi.nlm.nih.gov/15375829/)]
35. Unnanuntana A, Ruangsombon P, Keesukpant W. Validity and responsiveness of the two-minute walk test for measuring functional recovery after total knee arthroplasty. *J Arthroplasty* 2018 Jun;33(6):1737-1744 [FREE Full text] [doi: [10.1016/j.arth.2018.01.015](https://doi.org/10.1016/j.arth.2018.01.015)] [Medline: [29454556](https://pubmed.ncbi.nlm.nih.gov/29454556/)]
36. Rosenfeldt A, Penko A, Bazyk A, Streicher M, Dey T, Alberts J. The two minute walk test overground and on a self-paced treadmill detects dual task deficits in individuals with Parkinson's disease. *J Aging Phys Act* 2019 Dec 01;27(4):843-847 [FREE Full text] [doi: [10.1123/japa.2018-0264](https://doi.org/10.1123/japa.2018-0264)] [Medline: [31034314](https://pubmed.ncbi.nlm.nih.gov/31034314/)]
37. Chan W, Pin T. Reliability, validity and minimal detectable change of 2-minute walk test, 6-minute walk test and 10-meter walk test in frail older adults with dementia. *Exp Gerontol* 2019 Jan;115:9-18 [FREE Full text] [doi: [10.1016/j.exger.2018.11.001](https://doi.org/10.1016/j.exger.2018.11.001)] [Medline: [30423359](https://pubmed.ncbi.nlm.nih.gov/30423359/)]
38. Wang Q, Ye L, Luo H, Men A, Zhao F, Ou C. Pedestrian walking distance estimation based on smartphone mode recognition. *Remote Sensing* 2019 May 13;11(9):1140. [doi: [10.3390/rs11091140](https://doi.org/10.3390/rs11091140)]
39. Shih WY, Chen LY, Lan KC. Estimating walking distance with a smart phone. In: *Proceedings of the Fifth International Symposium on Parallel Architectures, Algorithms and Programming*. 2012 Presented at: Fifth International Symposium on Parallel Architectures, Algorithms and Programming; Dec. 17-20, 2012; Taipei, Taiwan p. 166-171. [doi: [10.1109/paap.2012.33](https://doi.org/10.1109/paap.2012.33)]

40. Barin L, Vaney C, Puhan M, von Wyl V. Recommended outcome measures for inpatient rehabilitation of multiple sclerosis are not appropriate for the patients with substantially impaired mobility. *Mult Scler Relat Disord* 2018 May;22:108-114 [[FREE Full text](#)] [doi: [10.1016/j.msard.2018.04.001](https://doi.org/10.1016/j.msard.2018.04.001)] [Medline: [29655044](https://pubmed.ncbi.nlm.nih.gov/29655044/)]
41. Scalzitti D, Harwood K, Maring J, Leach S, Ruckert E, Costello E. Validation of the 2-minute walk test with the 6-minute walk test and other functional measures in persons with multiple sclerosis. *Int J MS Care* 2018 Jul;20(4):158-163 [[FREE Full text](#)] [doi: [10.7224/1537-2073.2017-046](https://doi.org/10.7224/1537-2073.2017-046)] [Medline: [30150899](https://pubmed.ncbi.nlm.nih.gov/30150899/)]
42. Stellmann J, Neuhaus A, Götze N, Briken S, Lederer C, Schimpl M, et al. Ecological validity of walking capacity tests in multiple sclerosis. *PLoS One* 2015 Apr 16;10(4):e0123822 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0123822](https://doi.org/10.1371/journal.pone.0123822)] [Medline: [25879750](https://pubmed.ncbi.nlm.nih.gov/25879750/)]
43. Feys P, Bibby B, Romberg A, Santoyo C, Gebara B, de Noordhout BM, et al. Within-day variability on short and long walking tests in persons with multiple sclerosis. *J Neurol Sci* 2014 Mar 15;338(1-2):183-187 [[FREE Full text](#)] [doi: [10.1016/j.jns.2014.01.001](https://doi.org/10.1016/j.jns.2014.01.001)] [Medline: [24439144](https://pubmed.ncbi.nlm.nih.gov/24439144/)]
44. Gijbels D, Alders G, Van Hoof E, Charlier C, Roelants M, Broekmans T, et al. Predicting habitual walking performance in multiple sclerosis: relevance of capacity and self-report measures. *Mult Scler* 2010 May;16(5):618-626 [[FREE Full text](#)] [doi: [10.1177/1352458510361357](https://doi.org/10.1177/1352458510361357)] [Medline: [20207785](https://pubmed.ncbi.nlm.nih.gov/20207785/)]
45. Montalban X, Mulero P, Midaglia L, Graves J, Hauser SL, Julian L, et al. FLOODLIGHT: smartphone-based self-monitoring is accepted by patients and provides meaningful, continuous digital outcomes augmenting conventional in-clinic multiple sclerosis measures. In: Poster Presentation at the 34th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS). 2018 Presented at: 34th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS); October 10-12, 2018; Berlin, Germany p. 624 URL: <https://onlinelibrary.ectrims-congress.eu/ectrims/2018/ectrims-2018/228468/xavier.montalban.floodlight.smartphone-based.self-monitoring.is.accepted.by.html>
46. Ancker J, Witteman H, Hafeez B, Provencher T, Van de Graaf M, Wei E. "You Get Reminded You're a Sick Person": personal data tracking and patients with multiple chronic conditions. *J Med Internet Res* 2015 Aug 19;17(8):e202 [[FREE Full text](#)] [doi: [10.2196/jmir.4209](https://doi.org/10.2196/jmir.4209)] [Medline: [26290186](https://pubmed.ncbi.nlm.nih.gov/26290186/)]
47. Birkhoff S, Smeltzer S. 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 [[FREE Full text](#)] [doi: [10.1111/jnu.12298](https://doi.org/10.1111/jnu.12298)] [Medline: [28605151](https://pubmed.ncbi.nlm.nih.gov/28605151/)]

Abbreviations

- 2MWT:** 2-Minute Walking Test
- 6MWT:** 6-Minute Walking Test
- EDSS:** Expanded Disability Status Scale
- HC:** healthy control
- ICC:** intraclass correlation coefficient
- MICD:** minimally important clinical difference
- MS:** multiple sclerosis
- SDC:** smallest detectable change
- s2MWT:** smartphone-based 2-Minute Walking Test
- T25FW:** Timed 25-Foot Walk

Edited by G Eysenbach; submitted 28.03.21; peer-reviewed by M Cohen, V Tiberius; comments to author 27.04.21; revised version received 22.06.21; accepted 16.09.21; published 17.11.21.

Please cite as:

van Oirschot P, Heerings M, Wendrich K, den Teuling B, Dorssers F, van Ee R, Martens MB, Jongen PJ
A Two-Minute Walking Test With a Smartphone App for Persons With Multiple Sclerosis: Validation Study
JMIR Form Res 2021;5(11):e29128
URL: <https://formative.jmir.org/2021/11/e29128>
doi: [10.2196/29128](https://doi.org/10.2196/29128)
PMID: [34787581](https://pubmed.ncbi.nlm.nih.gov/34787581/)

©Pim van Oirschot, Marco Heerings, Karine Wendrich, Bram den Teuling, Frank Dorssers, René van Ee, Marijn Bart Martens, Peter Joseph Jongen. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 17.11.2021. 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 a Strengths-Based mHealth Tool (MyStrengths): Explorative Feasibility Trial

Stian Jessen^{1,2}, MPhilEd; Jelena Mirkovic¹, PhD; Elanor Halvorsen Brendmo¹, MA; Lise Solberg Nes^{1,3}, LP, PhD

¹Department of Digital Health Research, Division of Medicine, Oslo University Hospital HF, Oslo, Norway

²Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

³Department of Psychiatry and Psychology, Mayo Clinic, College of Medicine and Science, Rochester, MN, United States

Corresponding Author:

Stian Jessen, MPhilEd

Department of Digital Health Research

Division of Medicine

Oslo University Hospital HF

Aker Sykehus

Trondheimsveien 23

Oslo, 0586

Norway

Phone: 47 9804 9759

Email: stian.jessen@rr-research.no

Abstract

Background: As the number of people living with chronic illnesses increases, providing wide-reaching and easy-to-use support tools is becoming increasingly important. Supporting people in this group to recognize and use more of their personal strengths has the potential to improve their quality of life. With this in mind, we have developed the MyStrengths app prototype, a gamefully designed app aimed at aiding users in both identifying their strengths and using these strengths more actively in their daily life.

Objective: The goal of this study was to evaluate the user-reported feasibility and usefulness of the MyStrengths app. The study additionally aimed to explore whether the use of MyStrengths could be associated with selected psychosocial outcomes.

Methods: A 31-day explorative feasibility trial with a pretest-posttest design and an optional end of study interview was conducted. Data collection included system-use log data, demographic information, pre- and post-psychosocial measures (ie, strengths use, self-efficacy, health-related quality of life, depression), user experience measures (ie, usability, engagement, flow), and interview data.

Results: In total, 34 people with at least 1 chronic condition were enrolled in the study, with 26 participants (mean age 48 years, range 29-62 years; 1 male) completing the trial. Among these individuals, 18 were also interviewed posttrial. Participants used the MyStrengths app an average of 6 days during the trial period, with 54% (14/26) using the app over a period of at least 19 days. In total, 8738 unique app actions were registered. Of the psychosocial outcome measures, only 1 subscale, general health in the RAND 36-Item Health Survey, yielded significant pre- and posttest changes. Posttrial interviews showed that the number of participants who considered the MyStrengths app to be useful, somewhat useful, or not useful was evenly distributed across 3 groups. However, every participant did voice support for the strengths approach. All participants were able to identify a multitude of personal strengths using the MyStrengths app. Most participants that reported it to be useful had little or no previous experience with the personal strengths approach. A multitude of users welcomed the gameful design choices, particularly the rolling die feature, suggesting strengths exercises, activities that use a specific strength, were well received.

Conclusions: Although the reported usefulness and feedback from use varied, most participants were favorable to the strengths-focused approach to care and support. Consequently, low-threshold and wide-reaching mobile health tools that use a strengths-focused approach, such as MyStrengths, hold the potential to support people living with chronic illness in performing self-management and achieving mastery of their life.

(JMIR Form Res 2021;5(11):e30572) doi:[10.2196/30572](https://doi.org/10.2196/30572)

KEYWORDS

mHealth; personal strengths; gameful design; gamification; user engagement; explorative; feasibility; usefulness; usability; design; self-management; chronic illness

Introduction

Background

The number of people living with 1 or more chronic illnesses is increasing [1,2]. A recent review reported that up to 58% of people in developed countries live with multiple chronic diagnoses [3]. As the number of people living with chronic illness increases, finding new and more efficient ways of supporting them in living the best life possible with their illness is imperative. Self-management, or the activities, tasks, and skills one undertakes to manage life with illness, is one important aspect of such approaches [4]. Programs aiming to support and improve self-management often involve training and support that may aid in coping well with life with illnesses and include monitoring symptoms, adhering to medication regimens, and learning coping strategies [4,5].

Adding to these types of self-management approaches, research over the past few decades has also pointed to the potentially complementary benefits of exploring more “positive oriented” approaches to health care and nursing, such as focusing on people’s personal strengths when aiming to encourage, foster, and support self-management in chronic illness [6-9].

Personal Strengths

The concept of personal strengths has its foundation in positive psychology [7,10]. It has been defined as “traits/capabilities that are personally fulfilling, do not diminish others, are ubiquitous and valued across cultures, and aligned with numerous positive outcomes for oneself and others” [11]. Colloquially phrased, focusing on strengths means emphasizing what is valuable, possible, and doable, as opposed to the problem- and deficit-focused approach often present in modern medicine [12,13]. Being aware of and mobilizing one’s strengths may lead to a wide range of positive effects for people in the general population, such as well-being and better quality of life [14]. More specifically, Seligman et al [7] found related increases in happiness and decreases in depression, Proyer et al [15] found increases in well-being and happiness, Linley et al [16] found better goal accomplishment, and Lee et al [17] found increased resilience in the face of challenges. Similarly, Wood and colleagues [18] connected strengths use in general to well-being, vitality, self-esteem, positive affect, and reductions in stress.

Although much of the strengths-related research has included participants from the general population [11,19], previous studies from our research group [12,20,21] have identified a multitude of strengths important to people living with chronic illnesses. In addition to items common to general strengths classifications, such as the Values In Action strengths classification [11] (eg, being kind and caring, persistent, having a positive outlook on life, or having courage), people with health challenges have also reported strengths vital to them to include support from family and peers, positive relationships with health

care providers, and having helpful and constructive self-management strategies.

Strengths Activities and Exercises

Strengths are not a static part of an individual’s being but rather something malleable that can be changed and developed [11]. Nevertheless, although research related to activities that support people in recognizing and using their strengths, also called “strengths exercises,” have been published in the last few decades [11,22], few such studies and publications have focused on the potential of these strengths approaches for persons with chronic illnesses. Reviewing literature during the early phases of this project, our research team [23] identified 6 interventions described in 7 publications that included a personal strengths focus in chronic illness management and care [24-30]. The strengths activities used in these interventions included questionnaires asking participants to select strength(s) that apply the most to them [25,28,29], nominate their own strengths [26], use a specific strength in an activity each day [28,29] or to, in writing, reflect upon how they have used their strengths recently [24,26]. Although none of the studies reported on the development or adaption of the included strengths activities, the reported activities are still in line with common approaches for identifying and employing personal strengths [7,11,31-33]. These common approaches include using strengths in new and creative ways [7]; “strengths spotting,” which involves reflecting on how specific strengths are impacting you or trying to identify strengths in others [11,33]; the “three good things in life exercise” involving a daily routine of writing down three things that happened that day that you are grateful for and why [7,34]; and finally, performing “random acts of kindness” involving doing something that benefits others without directly benefiting oneself [35,36]. As with the outcomes of strengths use in general, as presented in the previous section, the identified studies employed a variety of outcome measures, including well-being [25-27,29,30], positive and negative affect [24,25,27], depression [24-26,29], strength-related outcomes such as self-esteem [27], and self-efficacy [30].

Strengths Support for People With Chronic Illness

Although “people living with chronic illness” reflect a large and heterogenous group, many of the challenges faced in daily life, such as fatigue, sleeping problems, difficult emotions, or energy loss, are shared across diagnoses [37]. Aiding people living with chronic illnesses to be aware of and use their personal strengths more may lead to better wellness outcomes [38], and interventions aiming to help increase people’s knowledge and use of their strengths may therefore be of great value for people living with chronic illnesses.

Both general or illness-specific patient educational activities and self-management training have been shown to benefit the patients with, for instance, improved self-management strategies and greater awareness of their condition [5], and to have positive health-economic impacts [39]. The majority of such programs appear to take place in municipalities and communities, while

approximately one-third are hospital based [5]. These programs are often group based and peer led, and the content varies, and may include providing general and diagnose-specific information, learning from and sharing with others in similar situations, or conducting exercise sessions [40]. These programs are also available in Norway, where this study was conducted, and are often in the form of lifestyle courses that seek to improve participants' quality of life by aiding them in finding unrecognized resources within themselves and strengthening their coping skills [5].

Still, not everyone living with chronic illness has the opportunity to participate in self-management training, and some have also reported feelings of loss or the need for follow-up sessions after the programs have ended [5]. Owing to the ubiquity of smartphones and other mobile devices, mobile health (mHealth) tools hold potential for reaching a much larger audience than do in-person interventions and other activities with care providers [41]. Given this, mHealth can add channels to existing services and also reach underserved populations or otherwise not active users of existing learning or educational programs [41].

Over the past few decades, an ever-increasing number of mHealth tools have been developed for various target groups, including weight [42] and stress management [43], medication adherence [44], smoking cessation [45], and oral hygiene [46]. Whereas traditional in-person health interventions have the benefits of health care providers or support personnel providing individual guidance, mHealth tools, as the name implies, are typically available to the user anywhere and anytime. Availability does not guarantee use, however, and these tools also need to be designed in ways that users find helpful, motivating, and engaging [47-49]. A popular approach to increase users' engagement with apps in general—and for mHealth specifically—has for the past decade been to create more playful and gamelike user experiences using design approaches and techniques known from the world of games [50-52]. Creating such gameful designs typically includes using features such as competitions, collaborations, narratives, or immersive visual designs to increase the users' value creation and enjoyment [50,53,54].

In sum, finding ways to provide and deliver strengths-focused support to people with chronic illnesses through personal digital devices could benefit users and widen the overall service reach in potentially cost-efficient ways.

The MyStrengths Project and App

Through earlier projects at our department, we have reported on multiple aspects related to the strengths approach to care and

self-management including on the conceptualization of strengths and health assets [38], patients insights into and requirements of the strengths-approach [21], and the strengths reported by patients [12,20]. With a high degree of participation from people in the user group, our research group created the MyStrengths app. The development process and design activities (eg, idea-generating and design workshops with people in the user group, seminars with an international advisory group of researchers, and workshops with game and health technology designers) have been described in detail in previous publications [23,55-57].

The MyStrengths app is designed to help its users identify and use more of their strengths. Its main feature is an assessment and subsequent overview of the users' strengths. Each of the strengths preprogrammed into MyStrengths—40 in total—is presented as a sphere floating on the home screen (Figure 1). The details regarding the development and identification of the 40 strengths chosen can be found in the work by Kristjánssdóttir et al [12]. Unassessed strengths are shown as empty green spheres that float up on the screen, a few at a time. When a strengths item is clicked on and opened (as seen in Figure 2), it can be read in full and rated as either yes (having the strength), partially (partially having the strength), or no (not having the strength); or the user can skip the strength. Depending on how the individual strengths are rated, they are colored red (having), yellow (partially), blue (not having), or green (skipped). Should the users feel that a strength is missing from the list, they can add additional strengths themselves. Having assessed ones' strengths, the app will suggest 2 to 3 strengths exercises for how to use and develop these (see Figure 3). When planned exercises are marked as completed, a short burst of celebratory stars showers the screen. The app also allows users to write a short reflection on the exercise and the strength used (see Figure 4).

Another key feature of the MyStrengths app is a logbook in which users can register how their day has been by choosing between 5 smiley faces ranging from sad to happy (see Figure 5). This logbook also asks the users to do the positive psychology exercise of entering three good things in life they experienced that day [7,58]. Through the presentation of all logged information, each day is summarized as cards with its “smiley status,” three good things, and strengths used (Figure 6).

To instill a sense of playfulness and surprise, the menu panel of the app (as seen at the bottom of Figures 1, 6, and 7) also has a die. When the die is pressed, a random exercise will be suggested (Figure 7) for one of the strengths that the user has rated as either having or partially having.

Figure 1. MyStrengths home screen.



Figure 2. Strengths assesment.

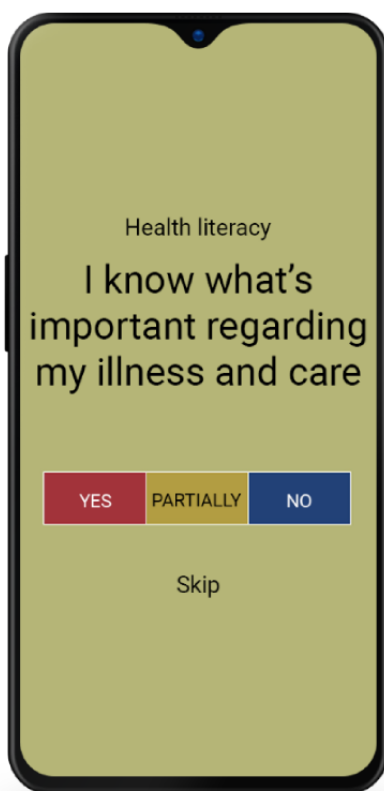


Figure 3. Exercise, with comment.

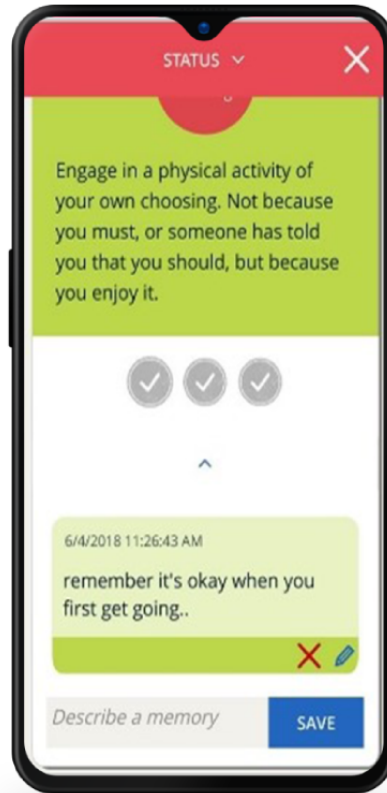


Figure 4. Reflections on strength.

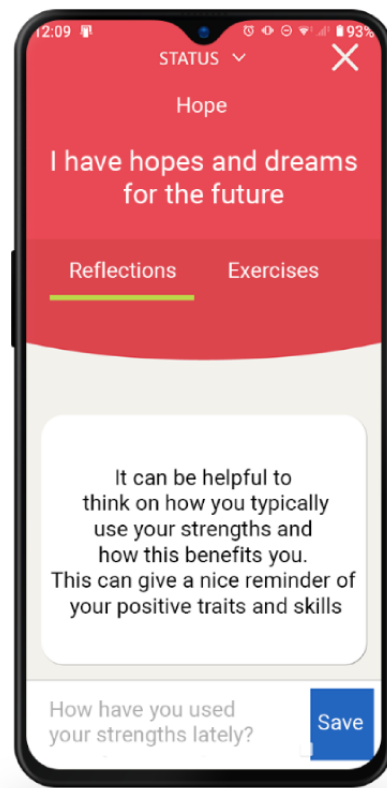


Figure 5. Smiley status and today's three good things.

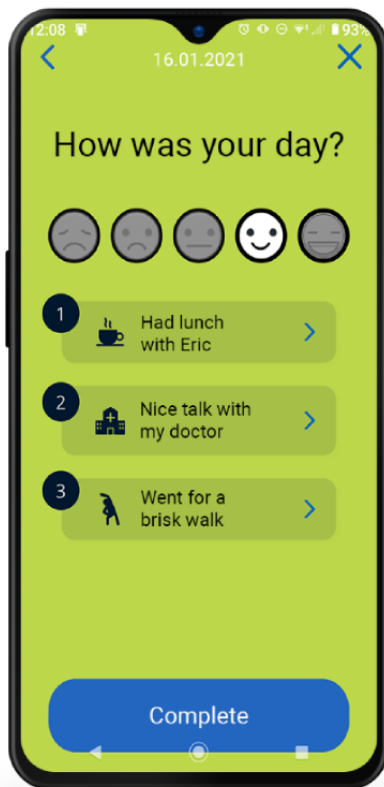


Figure 6. Log presentation.

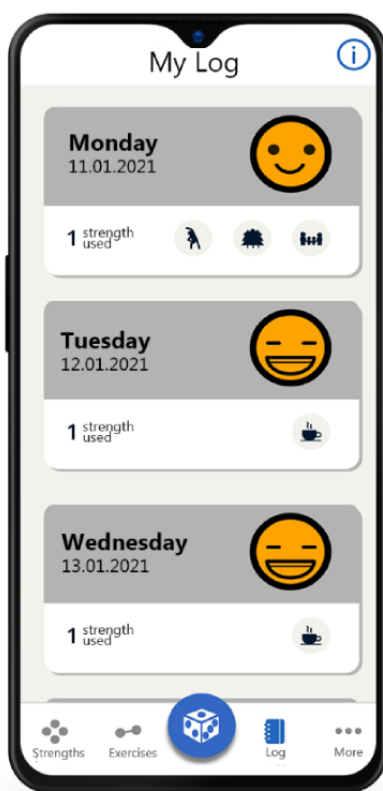
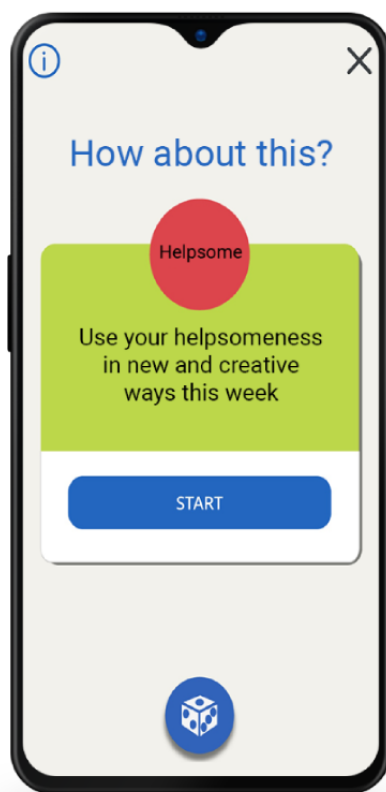


Figure 7. Die suggesting random exercise.



Aims

This study seeks to explore and evaluate ways of using mHealth to support people living with chronic illnesses in recognizing and using more of their personal strengths in their daily life to improve their overall well-being and quality of life. The study's primary aim is to, through a month-long explorative feasibility trial, examine the study participants' impression of and experiences with strengths-focused mHealth apps in general and the MyStrengths app specifically. Given the wide range of measures used to examine outcomes or efficacies of strengths-focused interventions for the target group [24-30], a secondary aim of this explorative feasibility trial is to explore which of these measures, if any, might preliminarily indicate effects from the use of the MyStrengths tool.

Methods

Recruitment

The project was planned and conducted in adherence to the principles of the Declaration of Helsinki [59] and approved by the Privacy Protection and Data Security committee at Oslo University Hospital (project #18/05449). All participants or their legal guardians signed informed consents before taking part in the study. Each participant received a gift card valued at Norwegian kr 250 (approximately US \$30) as compensation for participating.

Participants were recruited through multiple channels: through patient organizations or patient groups; through various health care providers, such as hospital educational centers and rehabilitation units; through advertising the trial on social media,

such as the research team webpage and Facebook; and through our departments existing network.

Criteria for inclusion were being over the age of 16 years and speak Norwegian, having 1 or more chronic illnesses (self-report), and being smartphone user (either Android version 4.6 or newer, or iOS version 11 or newer).

Study Design

Between November 2018 and March 2019, a 31-day feasibility trial was conducted using a mixed methods pretest-posttest design, consisting of outcome measures being completed before and after the trial period, logged data of participants' app use (ie, system use), and optional interviews posttrial. Using a mixed methods approach is well suited for investigating possible effects, or indications thereof, from using the MyStrengths app while also exploring the individual participants' experiences and views [60,61]. Participants enrolled at their convenience during the trial period, and later, after 14 days of use (halfway through the trial period), the third author (EHB) called each participant and checked whether they had any issues or questions regarding the app or study.

Data Types and Measures

To reach the study aims, a range of data types and measures were included.

Sociodemographic and Disease-Related Information

A study-specific questionnaire including questions related to age, gender, marital status, diagnosis, participant's time living with illness, and level of experience with smartphones and tablets was collected at baseline.

System-Use Log

Data logging each user's system use, with details related to app use, progress, and text input, were collected continuously, encrypted, and stored at the Services for Sensitive Data at the University of Oslo.

Psychosocial Outcome Measures

To achieve this study's secondary aim—to explore whether outcome measures used in strengths interventions [24–30] can indicate effects from the use of the MyStrengths app—we employed the following outcome measures as part of the pretest and posttest (ie, at baseline and after completion of the 31-day trial period).

Strengths Use Scale

The Strengths Use Scale [62] is a 14-item inventory measuring awareness and use of own personal strengths. The items are scored on a 7-point Likert scale from strongly disagree to strongly agree. Higher scores indicate higher use of one's strengths.

The Positive and Negative Affect Schedule

The Positive and Negative Affect Schedule [63] is a self-report questionnaire consisting of a 20-item scale describing the 2 main dimensions of people's mood: positive and negative affect. The items are scored on a 5-point Likert scale from very slightly or not at all to extremely. Higher positive affect indicates high energy and pleasurable engagement, and low positive affect indicates lethargy and sadness. High negative affect indicates distress or general unpleasurable engagements, whereas low negative affect indicates calmness and serenity [63].

Health-Related Quality of Life

The RAND-36 Measure of Health-Related Quality of Life [64] consists of 36 questions related to health-related quality of life, with 8 subscales (ie, physical functioning, role limitations caused by physical health problems, role limitations caused by emotional problems, social functioning, emotional well-being, energy/fatigue, pain, and general health perceptions). The response options are on various Likert scales (eg, 1-2, 1-5, 1-7). The scale options ranges include “all the time to not at all,” “absolutely right to absolutely wrong,” “nothing to strongly,” and “excellent to bad.” Higher scores indicate better functioning.

Self-Efficacy

The General Self-Efficacy Scale [65] assesses the strength of a person's belief in their ability to overcome challenges or respond to new and challenging situations. The scale consists of 10 statements, such as “I can solve most problems if I invest the necessary effort,” that are rated on a 4-point Likert scale, ranging from “not true at all” to “exactly true.” Higher scores indicate a higher degree of general self-efficacy.

The Center for Epidemiologic Studies Depression Scale

The Center for Epidemiologic Studies Depression Scale [66] measures depressive symptoms and contains 20 statements regarding the participants' feelings over the past week, such as “I cried” or “I felt happy.” These are rated on a 4-point Likert scale ranging from “never” to “almost the entire time,” with higher scores indicating the presence of more symptomatology.

User Experience Measures

Good usability, user-friendliness, and engagement are essential to the success and adoption of mHealth tools [67]. To explore the participants' experiences with the MyStrengths app, the study included 3 user experience measures in the posttrial test.

System Usability Scale

The System Usability Scale (SUS) [68] measures the users' perceived usability of MyStrengths. The SUS is a 10-item questionnaire used as an end-of-test subjective assessment of a system's overall usability [69]. Each item contains a statement (eg, “I found the system unnecessarily complex” or “I found the various functions in this system were well integrated”) that is scored on a 5-point Likert scale, ranging from “strongly disagree” to “strongly agree.” Higher scores indicate better usability.

Flow State Questionnaire

The “Flow State Questionnaire” of the Positive Psychology Lab [70] aims to measure the users' optimal experience or flow [67]; that is, the experienced absorption in the activity and whether the balance between challenges and skills is optimal. The questionnaire contains 20 items in the form of statements (eg, “I could effortlessly perform well” or “Time passed faster than I thought it did”) that are scored on a 5-point Likert scale ranging from “strongly agree” to “strongly disagree.” Higher scores indicate a higher level of flow.

Personal Involvement Inventory

Users' involvement with mHealth tools has been shown to be related to their intrinsic motivation for using the tools [67,71]. To measure this, we used the shortened version of the Personal Involvement Inventory [72], a 10-item self-report measure gauging involvement and engagement in a tool or service. It contains 10 different statements, all beginning with “To me, the MyStrengths app is:” that are rated on a 7-point Likert scale with bipolar adjectives at the extremes, including “important” versus “unimportant,” or “boring” versus “interesting.” Higher scores indicate more involvement.

Data Collection Procedures

The outcome measures and system-use log (app activity) were collected online through the secure Services for Sensitive Data at the University of Oslo, Norway. Participants received a link via email to access a secure web portal to access and complete the pre- and posttrial outcome measures. The pretest link was sent to participants as a baseline and completed before receiving access to the app. The posttest link was sent after the participants had completed the 31-day trial period. The system log data were encrypted by the app and then automatically sent to the secure Services for Sensitive Data.

After completing the posttest, all participants received a thank-you phone call from the research team, in which they were also invited to take part in a follow-up interview if interested.

The interviews were semistructured [73,74] and conducted by either the first (SJ) or third (EHB) author. Each interview lasted approximately 30 minutes and was guided by a semistructured interview guide covering 4 topics: (1) general impression(s),

(2) strengths functionalities, (3) design, and (4) additional input and free feedback. The complete interview guide, including suggested questions and follow-up questions, translated to English, is available in [Multimedia Appendix 1](#). To ensure the interviews were conducted similarly, the interviewers conducted the 2 first interviews together and subsequently made necessary adjustments to the interview guide. All interviews were audio-recorded, and the interviewer also took notes during the interviews.

Analysis

System-Use Log

The system use log was transferred from the secure Services for Sensitive Data and imported into Microsoft Excel [75] as one large comma-separated values file. During cleaning and sorting, only data from users who completed both the pre- and posttest were kept. This decision was made to allow us to have parity in the data from the use-log and quantitative outcome measures. Log data for 31 days (ie, from the first login) were used and examined for each participant. Using Power Queries in Microsoft Excel, the use metrics' final tallies were extracted.

Outcome Measures and Questionnaires

The outcome measures and questionnaires data were transferred from the secure Services for Sensitive Data and imported into Microsoft Excel as one spreadsheet. The data were cleaned and sorted by the first author (SJ), and, as with the system use log, only data from participants completing the trial were included. Only including data from users completing the trial allowed us to apply a repeated-measure design to analyze data, more easily identify any effects on the measures used, and facilitate having the same population for all the quantitative measures. The individual forms were imported into SPSS 26 (IBM Corp) [76], and the individual outcome measures' scores were calculated following each measure's specific guides and procedures. Pre- and posttest change scores were calculated, and their significance was tested using paired-sample *t* tests. Cohen *D* effect sizes [77] were also calculated for each measure.

Interviews

The interviews were analyzed using qualitative content analysis [78] following a selective coding strategy [73] that focused on the participants' use of and experiences with the MyStrengths app. The analysis was conducted by SJ and EHB, who also had conducted the interviews.

First, SJ and EHB read the transcribed interviews to familiarize themselves with the entire data corpus. The transcripts were imported into NVivo 12 (QSR International) [79], and, employing a directed approach [80] and using the main topics

from the interview guide (ie, general, strengths, functionalities, design, and free feedback) as initial codes, SJ and EHB separately coded 2 interviews. The codes were then discussed between the coders and merged into a new list used to code 4 more interviews. The code list was then discussed between SJ, JM, and EHB, and further adjustments were made. The updated code list was then used to code all interviews. At this stage, the coding process and code list were also discussed with a colleague well experienced with qualitative methods who served as an auditor. Disagreements were discussed until one final set of coded material was reached. In the end, the coding process resulted in 5 top-level codes: (1) background, (2) use of the app, (3) design and functionalities, (4) usability and user-friendliness, and (5) experienced usefulness. Further, the coding yielded 21 subcodes in level 2 and, below these, another 13 in level 3. Examples of the code hierarchy included design and functionality → strengths and assessment → spheres or colors (design), background → strengths → awareness, and usability → improvements → split up/hide/focus on strengths. Using the final code list, SJ and EHB authors coded the entire corpus. To illustrate the findings, quotes from participants are included throughout the presentation of the results. To ensure participants' anonymity, quotations presented are not linked with diagnosis or demographic information.

Results

Participants

In total, 34 participants were initially included in the trial. Of these, 2 withdrew from the study, 1 could not be contacted and thus did not respond postbaseline, 3 could not be contacted and thus did not respond after the follow-up call at approximately 2 weeks, 1 was unable to install the app and therefore did not participate in the trial, and 1 had technical issues with the posttest and therefore did not complete the follow-up questionnaire. Consequently, 26 participants completed the trial, including the pre- and postmeasures. After trial completion, the participants were invited to a posttrial interview, in which 18 agreed to participate. [Figure 8](#) presents the recruitment flow of the participants.

Demographic and background information for the 26 participants completing the feasibility trial is presented in [Table 1](#).

There were few demographic differences between the completers and noncompleters: 1 of the 8 noncompleters was male, and they had an average age of 54 (range 37-61); 7 of 8 had a lot of experience with smartphones or tablets; and 6 of the 8 reported 7 (maximum) fondness for mobile or computer technology.

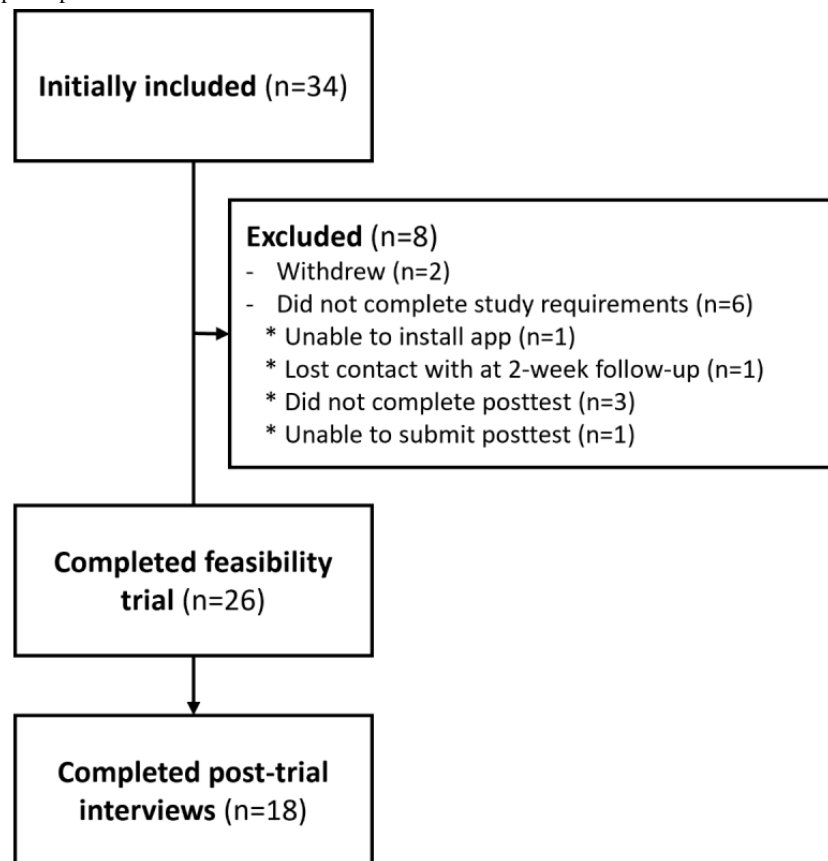
Figure 8. Recruitment and participant flow.

Table 1. Participants, characteristics, and demographic information (N=26).

Participant characteristic	Value
Age (years), mean (range)	46.8 (29-62)
Age (years), median	48
Gender, n (%)	
Men	1 (3.8)
Women	25 (96.2)
Marital status, n (%)	
Married/cohabiting	18 (69.2)
Single/divorced	8 (30.8)
Educational level, n (%)	
Elementary/high school	9 (34.6)
University < 4 years	14 (53.9)
University > 4 Years	3 (11.5)
Employment status, n (%)	
Full-time/part-time work	9 (34.6)
Student	2 (7.7)
Sick leave/disability benefits	15 (57.7)
Diagnosis^a, n (%)	
CFS ^b /ME ^c /fatigue/fibromyalgia	12 (46.2)
Mental health	5 (19.2)
Cancer/cancer related	3 (11.5)
Rheumatic	5 (19.2)
Chronic pain	8 (30.8)
Neurological (eg, MS ^d)	6 (23.1)
Other medical (eg, diabetes, kidney failure, Crohn)	8 (30.8)
Time living with illness (years), n (%)	
0-5	9 (34.6)
6-10	5 (19.2)
11-20	6 (23.1)
20+	5 (19.2)
Did not respond	1 (3.8)
Experience with smartphones/tablets, n (%)	
None	0 (0.0)
A little	1 (3.8)
Somewhat	6 (23.1)
A lot	19 (73.1)
Fondness for mobile/computer technology, n (%)	
1-2	0 (0.0)
3-4	1 (3.8)
5-6	10 (38.5)
7	15 (57.7)

^aParticipants could report having more than 1 diagnosis.

^bCFS: chronic fatigue syndrome.
^cME: myalgic encephalomyelitis.
^dMS: multiple sclerosis.

System-Use Log

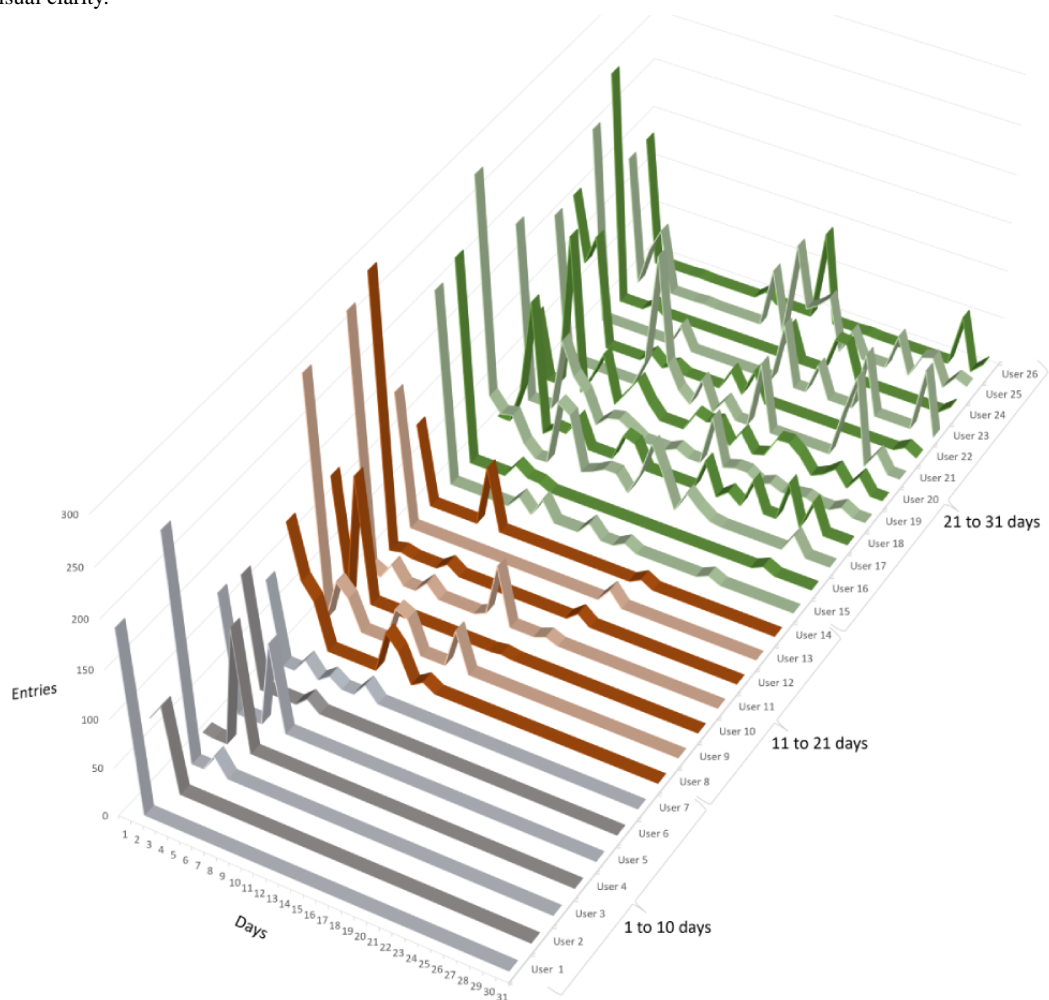
The 26 participants logged a total of 8738 unique actions during the 31-day trial period, and 14 of the 26 participants used the app for a period (days between first and last registered activity) of 19 days or more. Of these, 11 used it over a period of 26 days or longer. Table 2 shows the distribution of the participants' use

periods. On average, the users had 6 days with registered app activities. Figure 9 presents all 8738 logged data points from each user per day in the trial period (sorted after the last day of use). Although the users varied in both number of days with use and frequency, most activities took place early in the trial period, with as much as 45.65% (3989/8738) logged already on the first day of use.

Table 2. Use period (days between first and last registered app activity).

Period of use (days)	Participants, n (%), (N=26)
1	1/26 (4)
2-6	5/26 (19)
7-11	1/26 (4)
12-18	5/26 (19)
19-25	3/26 (12)
26-31	11/26 (42)

Figure 9. Data points per day per user. The x-axis displays each of the 31 days of the trial. The y-axis displays the number of registrations per day. The z-axis displays 1 lane each for the 26 users. The 3 colors are tiers of the last day with uses 1-10, 11-20, and 21-31, and with light and dark shades alternating for visual clarity.



The statistics of use for the unique features of the MyStrengths app are presented in Table 3. The participants, on average, rated 37 strengths as “having” or “partially having.” On average, only 8.6 strengths were rated as “not having.” Ten of the participants did not rate any strengths as “not having.” It should be noted that users could change the rating on their strengths, and thus some strengths are rated multiple times. Most of the participants (ie, 24) started at least 1 of the preprogrammed

strengths exercises, and half of these also added their own exercise(s). Sixteen participants added items to the “three good things in life” activity, which they did an average 8.9 times. The die, which suggests strengths exercises from strengths rated by users as “having” or “partially having,” was used by 23 participants, who in total started 86 suggested strengths exercises.

Table 3. Features and use statistics (N=26).

Feature	Participant use, n (%)	Entries		
		Number	Mean	Median
Strengths				
Strengths rated total	26 (100)	1122	43.2	37.5
Strengths rated as “Having”	26 (100)	584	22.5	20.0
Strengths rated as “Partially Having”	25 (96.2)	362	14.5	13.0
Strengths rated as “Not Having”	16 (61.5)	138	8.6	4.5
Strengths skipped/not rated	5 (19.2)	38	7.6	9.0
Strengths reflection added (free text)	17 (65.4)	101	5.9	3.0
Custom strengths added (free text)	2 (7.7)	3	1.5	1.5
Strengths exercises				
Exercises started	24 (92.3)	192	8.0	5.0
Exercise rated	19 (73.1)	139	7.3	5.0
Custom exercises added (free text)	12 (46.2)	44	3.7	3.0
Exercise comment added (free text)	15 (57.7)	84	5.6	3.0
Daily log				
Three good things in life added (free text)	16 (61.5)	143	8.9	5.0
Daily log smiley face registrations	21 (80.8)	100	4.8	3.0
Die				
Die menu opened	23 (88.5)	83	3.6	2.0
Suggested exercise started	20 (76.9)	86	4.3	2.0

Free-Text Entries

The MyStrengths app allows users to add comment or free text of up to 500 characters at various points in the app. The logging also recorded these. An overview of these types of entries is presented in the following sections.

Strengths and Strengths Reflections

In total, 17 of the 26 users added 101 free-text entries as reflections on their strengths. These entries ranged from single word affirmations that they had this strength (eg, “yes”) to summaries of activities or situations they had taken part in (eg, “played with the children today,” “getting better at this”). Some comments were celebratory and regretful at the same time. For instance, connected to the strength “I know how to set limitations for myself,” a user wrote, “I have gotten better at setting limits for myself, but many others do not understand this.” Some of the reflections also had deeper, personal content. For instance, as a reflection on their strengths, one person wrote the following:

I appreciate what is good in life. After my child died [as a teenager], life has morphed into something different. Still, I have a good life today, 5 years after.

A similar comment to the strength “I am at peace with my situation” came from another participant who wrote the following:

When my daughter died [some years ago], it was nice to be able to come to peace with the situation.

Although 17 of the participants added reflections to their strengths in the app, only 2 added a total of 3 new strengths into the app (ie, “I am open and share with others,” “I am efficient,” and “I am fond of animals”).

Strengths Exercises

Eighty-four free-text comments were added to the strengths exercises. Some were simple congratulatory comments participants included for themselves, such as “great” and “works OK, and is important,” or brief reflections, such as “good feeling both physically and mentally,” or “gives me strength and

confidence.” Some of the comments were also more comprehensive and contemplative, as exemplified by these two quotes:

Workout! Even if it's only taking a walk around the area. Time to myself is so important and shall be prioritized for my own, as well as my family's sake.

Participants added a total of 44 custom (or adapted) strengths exercises in the app. Most of these were simple and primarily specified activities the users planned to do, such as “take a walk,” “rest enough,” “meet NNN once a month,” “Yoga,” or “change medication dose.” As with the strengths reflections above, a few of these inputs were also of a more cognitive reflective nature, such as the following: “appreciate the little things in life: coffee, smells...”, “grow when facing obstacles, find people with knowledge, ask for help”, and “thank others for being good people in my life.”

Daily Log

In the daily log, 143 entries were added as “good things” of the day. Roughly half of these were brief, such as “worked 8 hours today” and “went for a walk.” The rest were more elaborate, for instance, the following:

Did a walk with dad, although I am not feeling great. Perfect weather!

as able to take breaks instead of pushing on. A bit drowsy and unwell today, so happy with that.

A few of the daily log comments were longer and reflective as well as personal, such as the following:

Been very tired the past few days, with lots of pain in the neck. Today started badly with a foul mood and arguing with the kids, and it gets worse for me when I am so frail. My daughter went off sad, returned, and

gave me a long hug before properly saying goodbye and leaving for school. That felt very good!<3

User Experience Measures

System Usability Scale

The system usability was rated as 61.5, which is considered a below-average experience. The center of the scale has been rated at 68, and the de facto industry standard for providing an above-average user experience is to achieve 80 points [81]. This points to the overall user-friendliness of MyStrengths being less than adequate.

Flow State Questionnaire

The main score was 2.83, which is slightly above the scale midpoint of 2.5. Considering the 2 subscales separately, the balance (task complexity) subscale averaged 2.58, and the absorption subscale averaged 2.90. This indicates that the overall flow intensity was medium, with the participants rating their engagement (absorption) with the tool as somewhat higher than that of the balance between task complexity and their skills.

Personal Involvement Inventory

The score was 3.85, which is slightly (0.35 points) above the scale's midpoint, pointing to an only slightly above medium involvement. This likely indicates that the participants did not find MyStrengths either fascinating, exciting, or appealing; or mundane, unexciting, or unappealing; but instead somewhere quite in between.

Psychosocial Outcome Measures

The pre- and posttrial outcome measures yielded effect sizes, Cohen *D* [77], mostly in the small to medium range. The paired-sampled *t* tests only yielded significant changes in the general health dimension of the RAND-36 Measure of Health-Related Quality of Life. The results from the statistical tests and calculations are presented in Table 4.

Table 4. Statistical results from the pre- and posttests.

Measure	Mean pretest (n=26)	Mean posttest (n=26)	Mean difference	SD	SE	t value	P value	Effect size (Cohen D)
Self-efficacy ^a	2.82	2.95	0.13	0.43	0.08	1.56	.13	0.31
PANAS^b								
Positive affect	24.12	23.65	-0.46	4.12	0.81	-0.57	.57	-0.11
Negative affect	22.31	22.23	-0.08	6.03	1.18	-0.07	.95	-0.01
CES-D ^c	21.12	19.12	-2.00	7.12	1.40	-1.43	.16	-0.28
Strengths use ^d	4.29	4.49	0.21	1.01	0.20	1.04	.31	0.20
Health-related quality of life^e								
Physical function	59.23	60.38	1.15	18.13	3.56	0.33	.75	0.06
Social function	41.35	41.35	0.00	13.69	2.69	0.00	.99	0.00
Role physical	12.50	19.23	6.73	26.98	5.29	1.27	.22	0.25
Role emotional	75.64	70.51	-5.13	45.89	9.00	-0.57	.57	-0.11
Mental	71.85	73.23	1.38	7.83	1.54	0.90	.38	0.18
Vitality	26.15	28.27	2.12	14.57	2.86	0.74	.47	0.15
Pain	36.73	37.36	0.63	21.86	4.29	0.15	.89	0.03
General health	31.15	37.31	6.15	14.02	2.75	2.24	.03	0.44

^aGeneral Self-Efficacy Scale.

^bPANAS: Positive and Negative Affect Schedule.

^cCES-D: Center for Epidemiologic Studies Depression scale.

^dGeneral Strengths Use Scale.

^eRAND-36 Measure of Health-Related Quality of Life.

Interviews

The interviews covered many facets of the use of the MyStrengths app as well as of life living with chronic illness in general, and data material coding and analysis identified 4 major topics: strengths and overall thoughts on MyStrengths, features and functionalities, design and usability, and the strengths approach. These main topics are presented in the following sections.

Strengths and Overall Thoughts on MyStrengths

Familiarity with Strengths

Most participants reported having some familiarity with thinking about their strengths, and only 2 participants stated that this was a new concept to them. When describing their connection to strengths, 1 participant, who had been ill for many years, described her strengths focus as a natural part of her daily life:

... That is what I feel I have been good at over the years, thinking of my strengths—and using them.

Several interviewees reported having become familiar with thinking of their strengths through courses and rehabilitation activities they had attended over the years. When comparing the MyStrengths app to previous experiences with strengths-thinking, one participant said the following:

Through the years, and this is the ninth time I am at physical rehabilitation, we have tried to focus on...

on these things. But, it has always been—you know, together with others, so it is really nice to be able to do it at home too.

Perceived Usefulness of MyStrengths

In terms of their perceived usefulness of the MyStrengths app, the interviewed participants were evenly split 3 ways.

First, 5 of the interviewed participants that completed the trial reported finding the MyStrengths app useful and described enjoying the app and its novelty. Some of these users were new to the strengths concepts and appreciated the newly raised awareness about strengths, while others reported already being aware of some of their strengths yet found the strengths exercises beneficial. In describing the benefits of the MyStrengths app, one user succinctly summarized her experience:

...I feel like I have become more aware of my strengths, and that I have used them more consciously, and this has already led to... well actually positive changes in my life... I thought I already knew this, but it is something about using it yourself and becoming more aware.

Five of the interviewed participants found the app to be somewhat useful, both liking and disliking various aspects of it and describing the design and the content to be adequate. One participant stated the following:

It was a couple of times when I kind of went “oh yeah, that is right, things are quite all right with those things, and those.” And I thought that was nice. But, it was not really revolutionary in any way.

Finally, 6 did not find the app useful. Reasons ranged from already being familiar with ones' strengths, favoring a more straightforward, to-the-point solution if in app form, or preferring the content be delivered as courses. Another repeated response from these users was that MyStrengths was “too comprehensive” and “demanding” due to many questions, prompts, or requirements for reflection.

Features and Functionalities Of MyStrengths

Strengths Assessment

The predefined list of strengths in the MyStrengths app contains 40 items. The participants, on average, rated more than half the strengths as “having,” and during the interviews, participants stated that seeing their strengths motivated them, gave them something to be proud of, or reminded them of, for instance, of the positive things in their life. No interviewees reported missing strengths outside the 40 included items, and most reported no issues with seeing strengths that did not apply to them, with several also describing it as virtually impossible to make a tool that fits every user perfectly. However, some did describe finding the number of predefined strengths somewhat overwhelming. Some also found it challenging to assess the 40 items in one session and often split the assessment over multiple days: “I split it over 2 days...., it would have been too much to answer all in one go.” Another user, who also thought the number of strengths was high, would have preferred to limit the number of strengths visible at one time in the app:

I thought there were quite a lot of them, yeah. And it kind of became too many for me to go into each of them and... think it over... actually. So, had it been fewer, it might have been easier for me to use.

Strengths Exercises

The majority of participants had started at least 1 strengths exercise in the app. Many of the participants interviewed described having enjoyed the concept of strengths exercises, and one participant voiced a common sentiment in describing these:

... Obviously, it feels a bit more committing to fulfill what you have written down... I mean, it feels like something you must do as, well, one has promised to do it either to oneself or... yeah...

Several of the participants wanted more strengths-exercises to, in the words of one participant, “make it more interesting for long term use.”

The users varied in their appreciation of the different types of exercises. On average, the interviewees seemed to prefer physically oriented strengths exercises (eg, “aim to do something new this week,” or “show someone you care about that they are important to you, tell them why”), over the more cognitively oriented ones (eg, “think of one of your goals and break it down into smaller actions” or “take a few minutes and consider what it means to you to live in a safe environment”). Some of the

approximately 80 suggested strengths exercises in the app were experienced as repetitions of things the participants described having done before in other settings or courses. A few also reported that the exercises could be either too time or resource intensive, and some described this part of the app as a bit overwhelming or too complex. Nonetheless, several interviewees also noted that the strengths exercises could help them better understand the connected strengths and their relevance to them.

Daily Log and Three Good Things Exercise

Sixteen of the participants added items to the “three good things” exercise, and many were already familiar with the exercise. Of the participants interviewed, everyone reported appreciating this exercise, particularly how it could help them get an overview of that day and emphasize their successes.

That is something I think is really nice, and I have good experiences with... to write down 3 good things and look back on that day, to see that I have actually mastered something. That is good to keep in mind and can be really uplifting and supportive.

Some additionally suggested having a more extensive range of icons to select from within this exercise and using prompts or notifications to remind them to do the registration daily or point out good things they had already registered.

Die feature

Most of the interviewed participants had used the die menu at least once, and many described the die both as a useful and engaging feature. These participants reported liking that it was a straightforward way to access the strengths exercises and that it would suggest strengths exercises they would not necessarily choose themselves. One participant, for example, likened the die to being challenged by other people:

I found it fun to press that [the die], and then see what would show up and whether it is something I could manage. Obviously, you could not do everything then and there. But it is nice to... well, I like to kind of, be challenged by someone, and not come up with everything myself, because then I tend to choose a bit easier and safer, I dare say, if I were to choose for myself.

A number of the interviewed participants also noted the engaging factor of using the die. Both in terms of how it provided surprises and how it was interesting to browse through the various strengths exercises presented.

I especially liked the dice... how you just can press it and then «ahh», maybe I will do that, and then I opt it out, maybe not that one... And there are so many small exercises or tasks. I think, I liked that there are so many suggestions...

Design and Usability

General Design

A number of those interviewed reported they liked the overall design of the MyStrengths app. Of the 18 interviewees, 9 generally liked the design, even though the degree of enthusiasm varied. For example, one participant stated the following:

Well, I think it was really nice. But... you know, nothing either “wow” or “non-wow.”

Another said the following:

I really like that [the MyStrengths app] quite well. It is... as we talked about, it seems a bit more playful, instead of a list and then bang, bang, bang—this is what you have to do. That is, like, very structured. I like that... with the spheres and that they float around; you kind of get the impression that it is not that serious but at the same time useful. And concerning, again back to mastery, you get a bit more pleasure out of working on these things.

Of the other half of the interviewed participants, 3 reported liking the design somewhat, and 6 reported not appreciating the app’s overall design, describing it as sometimes chaotic or cumbersome to use or sometimes being cognitively challenging.

Spheres

Approximately half of the interviewed participants reported liking the spheres, with one participant saying the following:

[It’s] a lot funnier, haha. Of course! I am like a small child who can press a button, and then things are colorful.

Some described the spheres as reminding them of a lava lamp, making the app appear somewhat “fresh” and calming. Others did, however, describe the spheres as somewhat confusing and difficult to track. Most participants reported appreciating the movement(s) of the strengths spheres, stating that movement made the app a bit more “alive.” At the same time, some participants also expressed some hesitance toward the movements, with one stating the following:

On a very tough day, it is tiring, but that is my personal opinion. Others might simply find it pleasant to see something with colors move slow, slowly...

Another user felt the same concern using the app on bad days but described being selective about which days she used MyStrengths, and stated the following:

I found it fun, haha... But, if I had a bad day, I might have gotten tired by it. Still, for the days I used it, it thought it was fine.

User-friendliness

The interviewed participants were also split related to the perceived user-friendliness of the MyStrengths app, with approximately half of the participants describing the app as user-friendly. For instance, these users remarked on how the colored spheres made them easy to distinguish by rating, how few steps were needed to access the strengths exercises (either through a specific strength or even easier by using the die), or that the limited set of features made it easy to use and navigate. Common issues raised by those who were less favorable were that the spheres changed places on the screen when moving between menus, they felt unsure as to what to do, and that it was easy to lose track of where they were in the app. Several participants suggested splitting the personal strengths assessment into multiple sessions or allowing the users to choose which spheres to keep active and onscreen.

Other comments related to use and usability included making MyStrengths available on larger devices such as tablets and PCs, using different colors on the spheres, or making the app more responsive. Lastly, some interviewees suggested the app could provide feedback on the actions done, such as congratulating them on completing an exercise, giving summaries of the strengths used in the past week, or, as mentioned earlier, adding options for notifications and reminders:

It had been super nice if you suddenly got a message which said “remember how good you are at this” kind of. Then I would go “maaw,” how nice... I would like that a lot.

The Strengths Approach

In discussing participants’ experience with MyStrengths, the interviews also broached topics that touch on the overall subject of strengths-based approaches.

Vulnerability and Ambivalence

Most participants reported that using their strengths more actively resulted in positive feelings such as mastery or confidence. As one participant stated the following:

I think it creates a positive awareness without... talking of that app, without it reminding me of... I am bad at doing this, in a way—and that you can only get better at it.

However, some of the participants described a sense of ambivalence concerning specific strengths. For instance, strengths or even strengths exercises that a person had or could do before becoming ill might now raise a sense of, or trigger, sadness. Also, some of the strengths listed in MyStrengths could be considered external, focusing on users’ surroundings, such as family members, care providers, or economy, something that the users may have little control over. Discussing this, one participant stated the following:

I did find some of the items [in the strengths assessment list] to be a bit, a bit touchy. For instance, the ones with “those [people] around you”; you know, it is about people close to you and how things are in your surroundings. And for instance, when you are ill for a long time, a lot of that gets lost... I mean, it is nice to look and reflect on those things, but at the same time, it is a bit... wow, cannot do anything about that. I would like to, but I cannot. Just does not work.

Strengths lost were not necessarily considered a negative thing, however, and some also considered these as regainable and reasonable goals for things to do in the future. When discussing working with strengths assessed as lacking, one participant said the following:

It is a bit mixed. But, obviously, there are strengths assessed as “not having” that one actually had before becoming ill and that you would rather have back. But, you also know that you will not manage that today. And then it’s okay to focus a bit on that. Even though you cannot do anything about it at the moment,

and you would have to have it as a goal to do something about it in the long run.

Exercising Strengths or Getting Rid Of Weaknesses?

When speaking of their strengths, many interviewees described these not only as positive attributes but as a concept existing in a dichotomous relationship with what they variably termed as weakness, faults, or errors.

Yeah, it was really that... or should I say, that you have somewhere you can go, to see how your day is working, or not, and to focus on yourself and your own stronger and weaker sides.

Most participants reported understanding and appreciating the concept of working on their strengths instead of their nonstrengths. Nevertheless, several reported also wishing to work on gaining strengths they perceived themselves as lacking. One participant, who was well experienced with mastery and strengths training, explained as follows:

I get happy when opening the app. It, it reminds me of what I manage and can. So yeah, there were new... I did actually get new tips through the exercises, things I was not aware of [...] What is nice with it is seeing all the strengths, that makes me happy, everything that's red [talking of the spheres rated as "having"], it makes you kind of want to go through the others [strengths], and get everything else red too [...]

However, not all participants sought to gain new strengths, and some chose to focus on the strengths they had, as one stated the following:

Well, it was very few spheres that did not turn red [marked as having] ... And of course, I feel I benefit from using and simply knowing that I am using, and being aware that when I am doing this, it is good for me. I am using some of my positive traits when doing this, and I think that is positive.

Future Use?

When asked about whether they would keep using the MyStrengths app after the trial, the participants were, again, roughly split into 3 groups, with 6 stating that they would, 4 stating they might, and 5 stating they would not continue using MyStrengths (3 interviewees did not answer this question). Reasons for wanting to continue using the app included finding benefits from using it, such as the app being a reminder of what they are able to achieve, emphasizing what they can manage in a situation when things do not always go as planned, creating awareness of their strengths without also making them think of the weaknesses, and facilitating a more deliberate use of their strengths which has led to positive changes. Several participants also noted how the app could motivate and support them to do tasks and activities that they would not necessarily be able to follow through with on their own:

When you can put those tasks and exercises into the app... To me, living alone, it feels like more of a commitment when I have added it as something I shall

do. It is motivating to know I have it there and that I can do something about it.

Among those unsure whether to keep using MyStrengths, several wanted to wait and see if new content or features, such as reminders and notifications, could be added. Participants could be put into 3 broad categories in terms of reasons for not wanting to keep using the app: (1) those who did not appreciate the overall design and visual concept, (2) those who did not find much use for the strengths exercises, and (3) those who reported already being quite familiar with their strengths. It should be noted, however, that even those in this last group found strengths that were new to them: "There was incredibly many there, and even some I did not even think of as strengths."

The Right Target Group

Regardless of what they thought about the MyStrengths app, all of the interviewed participants described appreciating the strengths approach, considered the strengths focus in chronic illness care to be an important one, and welcomed the MyStrengths app. One participant, a professional health care provider, would also recommend the app for people in her care:

I look forward to this, hopefully, being generally available for people with chronic illnesses so that I, as a professional, can recommend it. That would be really helpful.

Others reporting appreciating the MyStrengths app stated that it helped raise or increase awareness of their strengths. For instance, one participant stated the following:

Never done anything similar before. The first times I found it extremely hard... to see strengths... and after a while, you recognize it a bit more, that there are strengths, and you use them in this way... Sometimes, when not using the app, I think to myself, that is a strength you have used—and give myself a little pat on the shoulder.

However, several of those not describing the app as useful pointed out that even though they did not find it useful, the app could still benefit other groups of users. One interviewee suggested the following:

As I say, it is a very nice app, even though it did not really hit "bullseye" with me, and that some of the themes were a bit iffy... It is likely a much better fit for people who have not been "in the game" [of chronic illness] for as long as I have.

As such, it is possible that finding the right target group, rather than aiming for a "one-size-fits-all" approach, should be further explored.

Discussion

Principal Findings

Strengths interventions have been around for a long time, and as reported, many of the participants in this study had previously been introduced to strengths approaches through rehabilitation and educational courses. However, strengths-focused self-management support tools have not been available in the form of an app people can install on their phones. With

MyStrengths, users can have access to strengths support right there and then. All participants interviewed in this study voiced support for strengths-based tools and provided a range of positive and negative feedback related to MyStrengths' design and implementation. As there is little literature regarding strengths-focused mHealth tools, this trial has functioned both as an evaluation of the specific implementation, the MyStrengths app, and an exploration into the feasibility of strengths-focused mHealth interventions in general.

This study suggests that tools that focus on eliciting and mobilizing personal strengths, such as the MyStrengths app, may have the potential to support people living with chronic illness and to identify and use their strengths in their day-to-day life. The findings revealed 3 main points, which will be discussed in the following sections: (1) all participants could easily recognize multiple strengths within themselves; (2) the participants highly varied in their preferences for strengths exercises that they wish to practice; and (3) gameful designs can engage users, even when done in a soft and toned-down approach (ie, not emphasizing competitiveness or achievements).

Strengths Interventions

Everyone Has Strengths

On average, the participants rated more than half of the 40 strengths during the assessment in the MyStrengths app as "having" and under 9 as "not having." This is in congruence with one of the foundational underpinnings of the strengths concept: that everyone has strengths [7,13,82]. The findings also show that in a process similar to strengths nomination [11], the users were able to assess their strengths using the MyStrengths app.

Although the participants all reported a substantial number of strengths, the collected data did not identify the time it took users to find their first strength. Niemiec [11] has discussed the sometimes tricky task of getting users to recognize their first strength when new to strengths interventions. However, this may have a cascading effect, and identifying more strengths can subsequently become easier, thus highlighting the importance of making sure users can easily recognize their first strengths. During the development of the strengths list used in the MyStrengths app [12], prospective users contributed greatly to the process of nominating and formulating relatable strengths items. Based on the sheer number of strengths the individual users recognized from this list in the current study, participation of prospective users in these phases may be highly beneficial.

Some of the participants also highlighted feeling a sense of vulnerability when assessing strengths they experience as weakened or lost by changing life circumstances. Moreover, as the strengths assessment presents strengths at random, one might encounter situations where users start by rating many strengths as "not having" or are reminded of strengths that are now weakened or lost. Providing a positive onboarding experience that introduces and prepares the users for the rest of the app can be crucial, and starting this by not recognizing many strengths is likely counterproductive. Ways to optimize the start of the assessment can therefore be to make sure the first few items presented during strengths assessment, for instance, are typically

distinct (eg, being outgoing and social, having a supportive family, or being persistent), among the more popular strengths in the user group (ie, something that would need to be based on empirical evidence), or strengths that have been shown to be connected to life satisfaction (eg, love, gratitude, or zest) [83].

Internal and External Strengths

Typically, strengths tools and interventions are based around strengths that are internal to its users (eg, having zest, humility, or perseverance) [9,11]. However, MyStrengths also include external strengths and resources (ie, good relationships with health care providers, a supportive family, or living in a safe environment). The inclusion of these strengths stems from an earlier project by our research department [12], in which multiple external strengths were reported to be important by the participants. In the current study, however, several participants reacted negatively to some of these strengths, such as those related to caregivers or their economic situation. Some explained that these types of strengths are out of their control and are thus elements they cannot impact or work on acquiring. Addressing this type of issue, Kristjánsdóttir et al [84] grouped strengths into 4 categories (ie, personal strengths, strategies, resources in the environment, joy, and meaning), giving the users the choice of which category of strengths to assess. For MyStrengths, it might be that giving users the option to hide unwanted strengths could reduce the negative experiences from being reminded of strengths that are, for instance, out of reach.

Strengths or Weaknesses

Several interviewees reported wanting to work on what they considered their weaknesses, to regain strengths that had been weakened with their illness, or to simply gain as many strengths as possible instead of focusing on actually using their strengths. The benefits of the strengths approach come not only from knowing one's strengths or by having as many as possible but also from using them actively [11,13,83], and while working on one's strengths can lead to further growth, remedying one's deficiencies merely returns the person to a state of equilibrium [13]. Nonetheless, the wish for working on both strengths having and not having was often voiced by participants during the development of the MyStrengths app [23]. Therefore, the home screen was designed to give users an overview of all strength items rated, including those rated as "not having." To place primacy on the users' present strengths, however, the spheres are automatically sorted, and those rated as having are prioritized and shown above those either partially or not having. The broaden-and-build theory [85,86] posits that everyday positive emotions can contribute to a cascade of other positive emotions. Thus, if users can get a positive experience from mastering new things, such as working on strengths they do not (yet) possess, this might also contribute positively to the overall experience of the app. Additionally, allowing users to work on the strengths of their choice can also increase their sense of autonomy, contributing to motivation and overall positive experiences [87].

Strengths interventions are typically conducted face-to-face [11], affording the therapist or counselor to continuously adapt the session to the receiver. With mHealth tools, this type of guidance is usually not available, and making sure introductions and tutorials more than adequately cover the purpose and

rationale of the strengths-based approach should be of high importance. The presented findings indicate that providing just a short introduction when starting the app for the first time was insufficient. Therefore, more comprehensive approaches, such as giving users an in-person introduction where the app, its content, and rationale is presented and discussed in more detail may be beneficial [43].

Strengths Exercises

The feedback on the MyStrengths strengths exercises was mixed, and preferences for types of exercise seemingly depend on personal taste. Børøsund and colleagues [43] suggested that variation in exercises and exercise types is beneficial to support use. Similarly, some users in this study appeared to prefer exercises asking them to do something concrete, while others appreciated the more cognitively oriented exercises. A multitude of sources for strengths exercises and activities exist [7,11,33,38], yet little has been published on the actual development of such exercises and activities. Even though based on or strongly inspired by exercises described by existing literature, the exercises used in MyStrengths have been adapted to fit the connected strength. Although user evaluations have been conducted for these exercises [23], further improvement can be gained from the input provided by participants in our study. Additionally, the free-text inputs in the app during this trial yielded valuable insights into how users themselves write strengths exercises.

Based on the variety of individual tastes and preferences related to strengths, future strengths-focused tools should include multiple varied exercises or activities related to each strength. The pool of exercises could further be expanded by allowing users to share exercises they have created with other users. These two points could potentially increase the likelihood of users finding content to their liking and could also carry the additional benefit of users not having to repeat exercises but instead find new content for a more extended period of using the app.

The Right Target Group?

Regardless of what they thought about the MyStrengths app, all interviewed participants described appreciating the strengths approach and considered the strengths of chronic illness care to be important. However, those not finding the MyStrengths app useful did comment that it might be better suited for people relatively new to living with their health issues. This also reflects user feedback received throughout the development of MyStrengths [23], and future research should consider including more participants “new” to living with chronic illness when designing non-disease-specific mHealth support and educational tools.

Design and Usability

A Toned-down Gameful Design Approach

As mHealth tools naturally lack the person-to-person connection common when seeing a counselor or a physiotherapist, it is essential to design such tools in manners that increase the likelihood of them being used as intended [88]. With this in mind, gameful designs have become a common design approach

for mHealth apps specifically [50,54] as well as for apps in general [89]. Design features that are competitive or otherwise highlight the user’s performance, such as the number of strengths assessed or exercises completed, have nevertheless been avoided in the MyStrengths app. The reasoning for this approach is that, as per the strengths perspective, the focus should be on using one’s strengths, irrespective of how many there are, and not necessarily focusing on identifying and adding new ones [11,13]. Another critical aspect supporting this choice is that guidelines [90] and empirical studies [91] have cautioned against relying heavily on competitive elements in mental health and well-being interventions.

User engagement can also be encouraged through softer, more toned-down, gameful design approaches, such as creating user experiences that are visually pleasing and enjoyable [50]. Examples of this can include theming the app as a growing garden, in which growth is a metaphor for one’s progression and development [91,92] or using the metaphor of a journey toward a more flourishing life [49]. Using these kinds of metaphors was also suggested by participants in the co-design activities during the early development of MyStrengths [56]. The home screen of the MyStrengths app draws inspiration from the ubiquitous lava lamp (ie, a bottle-shaped lamp in which colored shapes of molten wax slowly floats around) and is designed to encourage engagement by presenting the users’ strengths in a playful and calming manner. Although interviewees in this study were split in their appreciation of this design concept, those in favor reported enjoying the colors and the movement of the spheres on the home screen.

One archetypical game design element that holds potential is randomness [93]. However, it has not been widely used within mHealth tools for well-being or mental health [94]. In MyStrengths, the die draws exercises at random, and many users reported liking this feature and used it to find exercises to their liking quickly. By throwing the die themselves, users choose to “ask” for exercises, which may also aid users in maintaining autonomy instead of the app unsolicited suggesting exercises. The die suggestions further seem to create similar experiences to, or function as a substitute for, social interaction. As voiced by several participants, the exercises suggested by the die were acted upon much in the same way as if they had come from other people. Thus, although the MyStrengths app does not include any social features, it still appears to have provided users with some form of social motivation, which can be very powerful [87,95]. Despite the seeming ubiquity of social features in mobile apps, none of the participants interviewed reported missing social features such as sharing, collaboration, or communication in MyStrengths.

Use and User-friendliness

The home screen is almost exclusively focused on the strengths and the assessment of strengths, while all other features are available after navigating to other screens in the app from the menu at the bottom of the screen (Figure 1). Sieverink et al [96] reported that users largely followed the structure and paths of use presented on the main screen. It might thus be that the focus on strengths and strengths assessment partially can explain the vast difference in use between the assessment and the other

features (see [Table 3](#)). For instance, 1122 strengths were rated, while only 192 exercises were started, or 143 “three good things in life” were added. On average, the participants used MyStrengths on 6 unique days during the trial period, and 45.65% (3989/8738) of the total logged activity was done on the first day of use alone. As presented in [Figure 9](#) users stopped using the app steadily throughout the trial period. A decline in use over time is, however, typical in eHealth and mHealth [97]. In their study of a mHealth tool for diabetes management, Adu and Colleagues [98] found that use steeply declined when reminders were disabled after 2 of the 3 weeks of the trial period. As also supported by participants during the interviews, it is likely that notifications or reminders from the MyStrengths app could thus have contributed to maintaining user engagement both in terms of duration and frequency of use.

When creating mHealth tools with gameful approaches, it is also essential to make sure such design choices do not make the tools overly complex or cumbersome to use [47,50]. In referencing the overall number of spheres on the home screen, many participants in this study found it overwhelming and complex to navigate and use, potentially particularly so, as one participant said, “on bad days.” Ease of use has been shown to be an important aspect of users' continued commitment to mHealth tools [99]. One feasible way of increasing the overall ease of use and user-friendliness could be to simplify the app by, for instance, allowing users to control the number of strengths spheres visible.

The literature on gameful mHealth design for well-being and mental health is still scarce [50,90]. In this study, ways of creating gameful designs without the use of proverbial game elements such as competitions or rewards have been presented. To help mature this field, future research should further explore approaches to engaging and gameful design in mHealth while also making sure these designs and features are thoroughly integrated into the tools in ways that are motivating, usable, and meaningful [49,90,100].

Future Steps

Further development and efficacy testing of the actual MyStrengths app depends on funding. However, the following steps are recommended for future related research.

Based on this study, we wish to highlight 3 aspects related to the use and usability of MyStrengths that could be improved. First, an option to reduce the number of strengths spheres visible on the screen at any given time and keep their place on the screen constant (ie, disabling the spheres floating) could be beneficial. Along the same line, a form of notifications should be included. This could allow feedback to the users, such as summaries of strengths used in the past week, and also encourage app use, including reminding users to do strengths-exercises or the “three good things in life” activity. Third, the number of exercises included could also be increased, including a variation to suit as many users as possible (eg, including both physical and cognitive or mental tasks).

Despite the nonsignificant changes identified from the psychosocial outcome measures, several of these (eg, the General Self-Efficacy Scale, the Center for Epidemiologic

Studies Depression Scale, and the General Strengths Use Scale) encouragingly yielded effect sizes in the small to medium range. Such effect sizes are not uncommon in psychosocial interventions (eg, Van Beugen et al [101]) and do not preclude potential clinical impact. For instance, in consideration of the possible reach of mHealth interventions, even small effect sizes can provide improvements to overall public health [102]. Although strengths use can impact people's well-being and quality of life [14], the exact mechanisms still are somewhat unclear, and thus future research should further explore these aspects. Also, based on feedback from participants in this study, future research should explore the possibility that strengths-focused tools might be best suited for people new to life living with illness or unfamiliar with the strengths approach.

Given these findings, an expansion of the demographics or background questionnaire to include previous experiences with strengths-focused tools and interventions as well as including a posttrial questionnaire gauging use and impressions of the tools, design, and content could also be of benefit. The interviews in the current study mostly covered these aspects, but a larger number of participants included in future studies might allow inferences to be drawn concerning what types or groups of users (eg, based on age, gender, diagnosis, familiarity with strengths, or time with illness) strengths-focused mHealth tools could be best suited for.

This study has, using the MyStrengths app as a vehicle, demonstrated the feasibility of strengths assessment and exercises in unguided mHealth tools, and we hope the knowledge gained and shared from this work can be employed in future internal (ie, departmental) and external projects designing and testing mHealth solutions.

Strengths and Limitations

This explorative feasibility study, combining interviews, outcome measures, and input from the system-use log, has some strengths and limitations. First, the gender balance of the participants in this study is heavily skewed, with 25 of 26 (96%) being women. Recruitment procedures strived to encourage more men to participate, unfortunately without success. Such imbalance of gender participation is not uncommon to mHealth research [103,104] and also mirrors the actual distribution of users of health care apps and digital services in Norway [105]. Second, as only users who had completed the 31-day trial were interviewed, we may lack feedback from users who potentially did not enjoy the app as much and therefore did not complete the trial.

On the other hand, it may be that these people already are thoroughly familiar with their strengths and could thus be considered outside the target group. Considering this, the participants interviewed in this study may even better represent the users. The heterogeneity of the participant group (eg, the variation in age, diagnosis, time with their illnesses) could also be considered a strength, as this allowed for a wide range of feedback and perspectives from the participants. Similarly, the fact that log data were only from participants completing the pre- and posttests (although the noncompleters shared the same basic background and demographical data) could be seen as a

limitation, as the log data in this case may be skewed toward those “liking” the app.

Strengths-focused mHealth intervention is a new field with little existing research. Given this, what mechanisms are in play and how to measure these are still not entirely clear. None of the included outcome measures in the current study could yield solid information about how such a strengths-focused mHealth intervention could potentially impact users living with chronic illness. However, being an explorative pilot study with a relatively low number of participants (26) from diverse backgrounds, the primary goal of this pilot study was not to gauge any possible effects, and the result should thus not be unexpected. This also means that although some of the measures used showed medium effect sizes, no firm conclusion can be drawn as to whether the measures used are sensitive to the effects of using strengths-focused and positively oriented mHealth tools. Future research should explore underlying mechanisms or effects stemming from strengths interventions further in more extensive and more robust studies.

There is a need for both more holistic and robust research on eHealth and mHealth tools [61,106], and despite the described limitations, the mixed methods approach [60] in this study, using 3 data types (ie, the system-use log, the outcome measures, and participant interviews), provides a solid empirical backing allowing this study to provide important knowledge regarding the feasibility of strengths-focused mHealth apps in general and the MyStrengths app specifically.

Conclusions

This study describes the explorative feasibility trial of the MyStrengths app, a mHealth tool that seeks to aid users in recognizing and employing their own strengths. MyStrengths is, to the best of our knowledge, the first mHealth tool of its kind. Through this feasibility trial, we have shown that all participants were able to identify a large number of their own strengths using the mHealth app. Although their preferences for exercises varied, most also found some they liked. The playful design elements (ie, the colored moving spheres and the die) were well received by parts of the participant group even though some also reported issues or limitations. There are also some indications that tools such as MyStrengths may be best suited for people relatively new to living with their illnesses. Based on this, future research should keep exploring the opportunities of both strengths-focused and gamefully designed mHealth apps.

Although impressions, reported usefulness, and feedback from use varied in this study, most participants reported being highly favorable to the strengths-focused approach to care and support. As such, mHealth tools such as MyStrengths may have the potential to support people living with chronic illness in a strengths-focused approach to self-management and mastery of their life. Further, as mHealth tools proverbially accompany users everywhere, creating strengths-focused mHealth tools may also increase the reach and availability of this form of support for those living with chronic illnesses.

Acknowledgments

Funding for this project was provided by the Norwegian Research Council (grant #248026).

The authors would like to thank their larger project team: Cornelia M Ruland, Ólöf Birna Kristjánsdóttir, Kurt Stange, Shirley M Moore, Luis Fernandez Luque, Hein De Vries, and Christina Bode.

Gratitude also goes to the staff at the learning and mastery centers at the Hospital of Southern Norway, Nordland Hospital, and Bærum Hospital, as well as the Youth Council at Akershus University Hospital and Frisklivssentralen Asker for help and coordination.

Lastly, thanks go out to all participants who took part throughout this project for their time, creativity, and effort.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide (translated from Norwegian to English).

[DOCX File, 25 KB - [formative_v5i11e30572_app1.docx](#)]

References

1. Global action plan for the prevention and control of NCDs 2013-2020. World Health Organization. 2013. URL: <https://www.who.int/publications/i/item/9789241506236> [accessed 2021-10-01]
2. Norwegian Institute of Public Health. Helsetilstanden i Norge 2018 (Health Status in Norway 2018). Norwegian Institute of Public Health. 2018. URL: <https://www.fhi.no/publ/2018/fhr-2018/> [accessed 2021-07-10]
3. Hajat C, Stein E. The global burden of multiple chronic conditions: A narrative review. *Prev Med Rep* 2018;12:284-293 [FREE Full text] [doi: [10.1016/j.pmedr.2018.10.008](https://doi.org/10.1016/j.pmedr.2018.10.008)] [Medline: [30406006](https://pubmed.ncbi.nlm.nih.gov/30406006/)]
4. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [doi: [10.1207/S15324796ABM2601_01](https://doi.org/10.1207/S15324796ABM2601_01)] [Medline: [12867348](https://pubmed.ncbi.nlm.nih.gov/12867348/)]

5. Stenberg U, Haaland-Øverby M, Fredriksen K, Westermann KF, Kvisvik T. A scoping review of the literature on benefits and challenges of participating in patient education programs aimed at promoting self-management for people living with chronic illness. *Patient Educ Couns* 2016 Nov;99(11):1759-1771. [doi: [10.1016/j.pec.2016.07.027](https://doi.org/10.1016/j.pec.2016.07.027)] [Medline: [27461944](https://pubmed.ncbi.nlm.nih.gov/27461944/)]
6. Niemiec R, Shogren K, Wehmeyer M. Character strengths and intellectual and developmental disability: A strengths-based approach from positive psychology. *Education and Training in Autism and Developmental Disabilities* 2017;52(1):13-25. [doi: [10.1007/978-3-319-59066-0_2](https://doi.org/10.1007/978-3-319-59066-0_2)]
7. Seligman MEP, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. *Am Psychol* 2005 Aug;60(5):410-421. [doi: [10.1037/0003-066X.60.5.410](https://doi.org/10.1037/0003-066X.60.5.410)] [Medline: [16045394](https://pubmed.ncbi.nlm.nih.gov/16045394/)]
8. Lopez SJ, Pedrotti JT, Snyder CR. *Positive Psychology - The Scientific and Practical Explorations of Human Strengths* (4th Edition). Thousand Oaks, California: SAGE Publications, Inc; 2019.
9. Swartz MK. A strength-based approach to care. *J Pediatr Health Care* 2017;31(1):1. [doi: [10.1016/j.pedhc.2016.10.008](https://doi.org/10.1016/j.pedhc.2016.10.008)] [Medline: [27939585](https://pubmed.ncbi.nlm.nih.gov/27939585/)]
10. Seligman MEP, Csikszentmihalyi M. Positive psychology: An introduction. *American Psychologist* 2000;55(1):5-14. [doi: [10.1037/0003-066x.55.1.5](https://doi.org/10.1037/0003-066x.55.1.5)]
11. Niemiec RM. *Character Strengths Interventions: A Field Guide for Practitioners*. Göttingen, Germany: Hogrefe Publishing; 2017.
12. Kristjansdottir OB, Stenberg U, Mirkovic J, Krogseth T, Ljoså TM, Stange KC, et al. Personal strengths reported by people with chronic illness: A qualitative study. *Health Expect* 2018 Feb 25:787-795. [doi: [10.1111/hex.12674](https://doi.org/10.1111/hex.12674)] [Medline: [29478260](https://pubmed.ncbi.nlm.nih.gov/29478260/)]
13. Rapp C, Goscha R. *The Strengths Model : A Recovery-Oriented Approach to Mental Health Services*. New York, United States: Oxford University Press; 2012:a.
14. Seligman MEP, Peterson C. Positive clinical psychology. In: Aspinwall LG, Staudinger UM, editors. *A Psychology of Human Strengths - Fundamental Questions and Future Directions for a Positive Psychology*. Washington, DC: American Psychological Association; 2003:305-317.
15. Proyer RT, Gander F, Wellenzohn S, Ruch W. Strengths-based positive psychology interventions: a randomized placebo-controlled online trial on long-term effects for a signature strengths- vs. a lesser strengths-intervention. *Front Psychol* 2015;456:1-14 [FREE Full text] [doi: [10.3389/fpsyg.2015.00456](https://doi.org/10.3389/fpsyg.2015.00456)] [Medline: [25954221](https://pubmed.ncbi.nlm.nih.gov/25954221/)]
16. Linley AP, Nielsen KM, Gillett R, Biswas-Diener R. Using signature strengths in pursuit of goals: effects on goal progress, need satisfaction, and well-being, and implications for coaching psychologists. *International Coaching Psychology Review* 2010;5(1):6-15. [doi: [10.1017/CBO9781107415324.004](https://doi.org/10.1017/CBO9781107415324.004)]
17. Lee JH, Nam SK, Kim A, Kim B, Lee MY, Lee SM. Resilience: a meta-analytic approach. *Journal of Counseling & Development* 2013 Jun 07;91(3):269-279. [doi: [10.1002/j.1556-6676.2013.00095.x](https://doi.org/10.1002/j.1556-6676.2013.00095.x)]
18. Wood A, Linley P, Maltby J, Kashdan T, Hurling R. Using personal and psychological strengths leads to increases in well-being over time: A longitudinal study and the development of the strengths use questionnaire. *Personality and Individual Differences* 2011 Jan;50(1):15-19. [doi: [10.1016/j.paid.2010.08.004](https://doi.org/10.1016/j.paid.2010.08.004)]
19. Park N, Peterson C, Szvarca D, Vander Molen RJ, Kim ES, Collon K. Positive Psychology and Physical Health: Research and Applications. *Am J Lifestyle Med* 2016;10(3):200-206 [FREE Full text] [doi: [10.1177/1559827614550277](https://doi.org/10.1177/1559827614550277)] [Medline: [30202275](https://pubmed.ncbi.nlm.nih.gov/30202275/)]
20. Rotegård AK, Fagermoen MS, Ruland CM. Cancer patients' experiences of their personal strengths through illness and recovery. *Cancer Nurs* 2012;35(1):E8-E17. [doi: [10.1097/NCC.0b013e3182116497](https://doi.org/10.1097/NCC.0b013e3182116497)] [Medline: [21558850](https://pubmed.ncbi.nlm.nih.gov/21558850/)]
21. Mirkovic J, Kristjansdottir OB, Stenberg U, Krogseth T, Stange KC, Ruland CM. Patient insights into the design of technology to support a strengths-based approach to health care. *JMIR Res Protoc* 2016 Aug 24;5(3):e175 [FREE Full text] [doi: [10.2196/resprot.5906](https://doi.org/10.2196/resprot.5906)] [Medline: [27558951](https://pubmed.ncbi.nlm.nih.gov/27558951/)]
22. Gottlieb LN. Strengths-based nursing. *Am J Nurs* 2014 Aug;114(8):24-32; quiz 33,46. [doi: [10.1097/01.NAJ.0000453039.70629.e2](https://doi.org/10.1097/01.NAJ.0000453039.70629.e2)] [Medline: [25036663](https://pubmed.ncbi.nlm.nih.gov/25036663/)]
23. Jessen S, Mirkovic J, Nes LS. MyStrengths, a Strengths-Focused Mobile Health Tool: Participatory Design and Development. *JMIR Form Res* 2020 Jul 24;4(7):e18049 [FREE Full text] [doi: [10.2196/18049](https://doi.org/10.2196/18049)] [Medline: [32706651](https://pubmed.ncbi.nlm.nih.gov/32706651/)]
24. Cohn MA, Pietrucha ME, Saslow LR, Hult JR, Moskowitz JT. An online positive affect skills intervention reduces depression in adults with type 2 diabetes. *J Posit Psychol* 2014;9(6):523-534 [FREE Full text] [doi: [10.1080/17439760.2014.920410](https://doi.org/10.1080/17439760.2014.920410)] [Medline: [25214877](https://pubmed.ncbi.nlm.nih.gov/25214877/)]
25. Müller R, Gertz KJ, Molton IR, Terrill AL, Bombardier CH, Ehde DM, et al. Effects of a tailored positive psychology intervention on well-being and pain in individuals with chronic pain and a physical disability: a feasibility trial. *Clin J Pain* 2016 Jan;32(1):32-44. [doi: [10.1097/AJP.0000000000000225](https://doi.org/10.1097/AJP.0000000000000225)] [Medline: [25724020](https://pubmed.ncbi.nlm.nih.gov/25724020/)]
26. Cheung EO, Cohn MA, Dunn LB, Melisko ME, Morgan S, Penedo FJ, et al. A randomized pilot trial of a positive affect skill intervention (lessons in linking affect and coping) for women with metastatic breast cancer. *Psychooncology* 2017 Dec 16;26(12):2101-2108 [FREE Full text] [doi: [10.1002/pon.4312](https://doi.org/10.1002/pon.4312)] [Medline: [27862646](https://pubmed.ncbi.nlm.nih.gov/27862646/)]
27. Cerezo MV, Ortiz-Tallo M, Cardenal V, De La Torre-Luque A. Positive psychology group intervention for breast cancer patients: a randomised trial. *Psychol Rep* 2014;115(1):44-64. [doi: [10.2466/15.20.PR0.115c17z7](https://doi.org/10.2466/15.20.PR0.115c17z7)] [Medline: [25153949](https://pubmed.ncbi.nlm.nih.gov/25153949/)]

28. Nikrahan GR, Laferton JAC, Asgari K, Kalantari M, Abedi MR, Etesampour A, et al. Effects of Positive Psychology Interventions on Risk Biomarkers in Coronary Patients: A Randomized, Wait-List Controlled Pilot Trial. *Psychosomatics* 2016;57(4):359-368 [FREE Full text] [doi: [10.1016/j.psych.2016.02.007](https://doi.org/10.1016/j.psych.2016.02.007)] [Medline: [27129358](https://pubmed.ncbi.nlm.nih.gov/27129358/)]
29. Nikrahan GR, Suarez L, Asgari K, Beach SR, Celano CM, Kalantari M, et al. Positive psychology interventions for patients with heart disease: a preliminary randomized trial. *Psychosomatics* 2016;57(4):348-358 [FREE Full text] [doi: [10.1016/j.psych.2016.03.003](https://doi.org/10.1016/j.psych.2016.03.003)] [Medline: [27137709](https://pubmed.ncbi.nlm.nih.gov/27137709/)]
30. Zangi HA, Mowinckel P, Finset A, Eriksson LR, Høystad, Lunde AK, et al. A mindfulness-based group intervention to reduce psychological distress and fatigue in patients with inflammatory rheumatic joint diseases: a randomised controlled trial. *Ann Rheum Dis* 2012 Jun;71(6):911-917. [doi: [10.1136/annrheumdis-2011-200351](https://doi.org/10.1136/annrheumdis-2011-200351)] [Medline: [22186709](https://pubmed.ncbi.nlm.nih.gov/22186709/)]
31. Linley A. *Average to A+: Realising Strengths in Yourself and Others*. Coventry, UK: CAPP Press; 2008:9781906366001.
32. Schutte N, Malouff J. The impact of signature character strengths interventions: a meta-analysis. *J Happiness Stud* 2018 May 21;20(4):1179-1196. [doi: [10.1007/s10902-018-9990-2](https://doi.org/10.1007/s10902-018-9990-2)]
33. Niemiec R, Wedding D. *Positive Psychology at the Movies: Using Films to Build Virtues and Character Strengths* (2nd Edition). Göttingen, Germany: Hogrefe Publishing; 2014.
34. Gander F, Proyer RT, Ruch W, Wyss T. Strength-based positive interventions: further evidence for their potential in enhancing well-being and alleviating depression. *J Happiness Stud* 2012 Sep 4;14(4):1241-1259. [doi: [10.1007/s10902-012-9380-0](https://doi.org/10.1007/s10902-012-9380-0)]
35. Pressman SD, Kraft TL, Cross MP. It's good to do good and receive good: The impact of a 'pay it forward' style kindness intervention on giver and receiver well-being. *The Journal of Positive Psychology* 2014 Oct 16;10(4):293-302. [doi: [10.1080/17439760.2014.965269](https://doi.org/10.1080/17439760.2014.965269)]
36. Pressman SD, Cohen S. Does positive affect influence health? *Psychol Bull* 2005 Nov;131(6):925-971. [doi: [10.1037/0033-2909.131.6.925](https://doi.org/10.1037/0033-2909.131.6.925)] [Medline: [16351329](https://pubmed.ncbi.nlm.nih.gov/16351329/)]
37. Lorig K, Holman H, Sobel D, Laurent D, González V, Minor M. *Living a Healthy Life with Chronic Conditions: Self-Management of Heart Disease, Arthritis, Diabetes, Asthma, Bronchitis, Emphysema and Other Physical and Mental Health Conditions* (4th Edition) 4th. Boulder, CO: Bull Publishing Company; 2012.
38. Rotegård AK, Moore SM, Fagermoen MS, Ruland CM. Health assets: a concept analysis. *Int J Nurs Stud* 2010 Apr;47(4):513-525. [doi: [10.1016/j.ijnurstu.2009.09.005](https://doi.org/10.1016/j.ijnurstu.2009.09.005)] [Medline: [19819452](https://pubmed.ncbi.nlm.nih.gov/19819452/)]
39. Stenberg U, Vågan A, Flink M, Lynggaard V, Fredriksen K, Westermann KF, et al. Health economic evaluations of patient education interventions a scoping review of the literature. *Patient Educ Couns* 2018 Jun;101(6):1006-1035. [doi: [10.1016/j.pec.2018.01.006](https://doi.org/10.1016/j.pec.2018.01.006)] [Medline: [29402571](https://pubmed.ncbi.nlm.nih.gov/29402571/)]
40. Stenberg U, Haaland Øverby M, Fredriksen K, Kvisvik T, Westermann KF, Vågan A, et al. Utbytte av lærings- og mestringstilbud. *Sykepleien* 2017(4):52-55. [doi: [10.4220/sykepleiens.2017.61278](https://doi.org/10.4220/sykepleiens.2017.61278)]
41. World Health Organization. *Global Diffusion of eHealth: Making Universal Health Coverage Achievable: Report of the Third Global Survey on eHealth*. Geneva: World Health Organization; 2016.
42. Willey S, Walsh JK. Outcomes of a Mobile Health Coaching Platform: 12-Week Results of a Single-Arm Longitudinal Study. *JMIR Mhealth Uhealth* 2016 Jan 08;4(1):e3 [FREE Full text] [doi: [10.2196/mhealth.4933](https://doi.org/10.2196/mhealth.4933)] [Medline: [26747611](https://pubmed.ncbi.nlm.nih.gov/26747611/)]
43. Børøsund E, Varsi C, Clark MM, Ehlers SL, Andrykowski MA, Sleveland HRS, et al. Pilot testing an app-based stress management intervention for cancer survivors. *Transl Behav Med* 2020 Aug 07;10(3):770-780 [FREE Full text] [doi: [10.1093/tbm/ibz062](https://doi.org/10.1093/tbm/ibz062)] [Medline: [31330023](https://pubmed.ncbi.nlm.nih.gov/31330023/)]
44. Pouls BPH, Vriezেকolk JE, Bekker CL, Linn AJ, van Onzenoort HAW, Vervloet M, et al. Effect of Interactive eHealth Interventions on Improving Medication Adherence in Adults With Long-Term Medication: Systematic Review. *J Med Internet Res* 2021 Jan 08;23(1):e18901 [FREE Full text] [doi: [10.2196/18901](https://doi.org/10.2196/18901)] [Medline: [33416501](https://pubmed.ncbi.nlm.nih.gov/33416501/)]
45. Pløhn T, Aalberg T. Using gamification to motivate smoking cessation. 2015 Presented at: European Conference on Games Based Learning; Oct 8-9, 2015; Steinkjer, Norway p. 431-438.
46. Fijačko N, Gosak L, Cilar L, Novšak A, Creber RM, Skok P, et al. The Effects of Gamification and Oral Self-Care on Oral Hygiene in Children: Systematic Search in App Stores and Evaluation of Apps. *JMIR Mhealth Uhealth* 2020 Jul 08;8(7):e16365 [FREE Full text] [doi: [10.2196/16365](https://doi.org/10.2196/16365)] [Medline: [32673235](https://pubmed.ncbi.nlm.nih.gov/32673235/)]
47. Gemert-pijnen LV, Kelders S, Jong N, Oinas-kukkonen H. Persuasive health technologies. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. *eHealth Research, Theory and Development - A Multidisciplinary Approach*. London & New York: Routledge; 2018:228-243.
48. van GJEW, Nijland N, van LM, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](https://doi.org/10.2196/jmir.1672)] [Medline: [22155738](https://pubmed.ncbi.nlm.nih.gov/22155738/)]
49. Ludden G, Kelders S, Snippert B. This is your life! The design of a positive psychology intervention using metaphor to motivate. 2014 Presented at: 9th International Conference on Persuasive Technology (PERSUASIVE 2014); May 21-23, 2014; Padua, Italy p. 179-190. [doi: [10.1007/978-3-319-07127-5_16](https://doi.org/10.1007/978-3-319-07127-5_16)]
50. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: A systematic review of the literature. *Internet Interv* 2016 Dec;6:89-106 [FREE Full text] [doi: [10.1016/j.invent.2016.10.002](https://doi.org/10.1016/j.invent.2016.10.002)] [Medline: [30135818](https://pubmed.ncbi.nlm.nih.gov/30135818/)]

51. Orji R, Moffatt K. Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health Informatics J* 2018 Mar;24(1):66-91 [FREE Full text] [doi: [10.1177/1460458216650979](https://doi.org/10.1177/1460458216650979)] [Medline: [27245673](https://pubmed.ncbi.nlm.nih.gov/27245673/)]
52. Hamari J, Koivisto J, Pakkanen T. Do persuasive technologies persuade? - A review of empirical studies. 2014 Presented at: PERSUASIVE 2014, The 9th International Conference on Persuasive Technology; May 21-23, 2014; Padua, Italy. [doi: [10.1007/978-3-319-07127-5_11](https://doi.org/10.1007/978-3-319-07127-5_11)]
53. Huotari K, Hamari J. A definition for gamification: anchoring gamification in the service marketing literature. *Electron Markets* 2016 Jan 15;27(1):21-31. [doi: [10.1007/s12525-015-0212-z](https://doi.org/10.1007/s12525-015-0212-z)]
54. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. *J Biomed Inform* 2017 Jul;71:31-48 [FREE Full text] [doi: [10.1016/j.jbi.2017.05.011](https://doi.org/10.1016/j.jbi.2017.05.011)] [Medline: [28536062](https://pubmed.ncbi.nlm.nih.gov/28536062/)]
55. Mirkovic J, Jessen S, Kristjansdottir OB, Krogseth T, Koricho AT, Ruland CM. Developing Technology to Mobilize Personal Strengths in People with Chronic Illness: Positive Codesign Approach. *JMIR Form Res* 2018 Jun 05;2(1):e10774 [FREE Full text] [doi: [10.2196/10774](https://doi.org/10.2196/10774)] [Medline: [30684404](https://pubmed.ncbi.nlm.nih.gov/30684404/)]
56. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. *JMIR Mhealth Uhealth* 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: [10.2196/11579](https://doi.org/10.2196/11579)] [Medline: [30552080](https://pubmed.ncbi.nlm.nih.gov/30552080/)]
57. Jessen S, Mirkovic J, Westeng M. Game experience preferences of people with chronic illnesses. New York, USA: ACM Press; 2018 Presented at: 10th Nordic Conference on Human-Computer Interaction (NordCHI 2018); Oct 1-3, 2018; Oslo, Norway p. 898-903. [doi: [10.1145/3240167.3240237](https://doi.org/10.1145/3240167.3240237)]
58. Mongrain M, Anselmo-Matthews T. Do positive psychology exercises work? A replication of Seligman et al. (2005). *J Clin Psychol* 2012 Apr;68(4):382-389. [doi: [10.1002/jclp.21839](https://doi.org/10.1002/jclp.21839)] [Medline: [24469930](https://pubmed.ncbi.nlm.nih.gov/24469930/)]
59. Ethical principles for medical research involving human subjects. World Medical Association. 2018. URL: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> [accessed 2021-07-10]
60. Creswell JW, Creswell JD. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th Edition). Thousand Oaks, CA: SAGE Publications Inc; 2018.
61. Sieverink F, Köhle N, Cheung K, Roefs A, Trompeter H, Keizer J. Evaluating eHealth. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. *eHealth Research, Theory and Development - A Multidisciplinary Approach*. London & New York: Routledge; 2018:290-318.
62. Govindji R, Linley P. Strengths use, self-concordance and well-being: Implications for strengths coaching and coaching psychologists. *International Coaching Psychology Review* 2007;2(2):1038. [doi: [10.1037/t01038-000](https://doi.org/10.1037/t01038-000)]
63. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988 Jun;54(6):1063-1070. [Medline: [3397865](https://pubmed.ncbi.nlm.nih.gov/3397865/)]
64. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health Econ* 1993 Oct;2(3):217-227. [doi: [10.1002/hec.4730020305](https://doi.org/10.1002/hec.4730020305)] [Medline: [8275167](https://pubmed.ncbi.nlm.nih.gov/8275167/)]
65. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*. Windsor, UK: NFER-NELSON; 1995:35-37.
66. Radloff LS. The CES-D Scale. *Applied Psychological Measurement* 2016 Jul 26;1(3):385-401. [doi: [10.1177/014662167700100306](https://doi.org/10.1177/014662167700100306)]
67. Triberti S, Kelders S, Gaggioli A. User engagement. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. *eHealth Research, Theory and Development - A Multidisciplinary Approach*. London & New York: Routledge; 2018:271-289.
68. Brooke J. SUS - A quick and dirty usability scale. *Usability Evaluation in Industry* 1996;189(194):4-7.
69. Sauro J, Lewis J. *Quantifying the User Experience Practical Statistics for User Research* (2nd Edition). Cambridge, MA: Elsevier Inc; 2016.
70. Magyaródi T, Nagy H, Soltész P, Mózes T, Oláh A. Psychometric properties of a newly established flow state questionnaire. *The Journal of Happiness & Well-Being* 2013;1(12):89-100. [doi: [10.4324/9780203071878-10](https://doi.org/10.4324/9780203071878-10)]
71. Kelders S. Involvement as a working mechanism for persuasive technology. : MacTavish T, Basapur S. editors. 10th International Conference, PERSUASIVE 2015 Chicago, USA; 2015 Presented at: 10th International Conference, PERSUASIVE 2015; June 3-5, 2015; Chicago, IL p. 3-14. [doi: [10.1007/978-3-319-20306-5_1](https://doi.org/10.1007/978-3-319-20306-5_1)]
72. Zaichkowsky JL. The personal involvement inventory: reduction, revision, and application to advertising. *Journal of Advertising* 1994 Dec;23(4):59-70. [doi: [10.1080/00913367.1943.10673459](https://doi.org/10.1080/00913367.1943.10673459)]
73. Braun V, Clarke V. *Successful Qualitative Research a Practical Guide for Beginners*. London, UK: SAGE Publications Ltd; 2013.
74. Kvale S. *Doing Interviews*. London, UK: SAGE Publications Inc; 2007.
75. Microsoft Excel. Microsoft Corporation. URL: <https://office.microsoft.com/excel> [accessed 2021-02-10]
76. IBM SPSS Statistics for Windows, Version 27.0. IBM Corp. 2020. URL: <https://www.ibm.com/analytics/spss-statistics-software> [accessed 2021-02-10]
77. Cohen J. A power primer. *Psychol Bull* 1992 Jul;112(1):155-159. [doi: [10.1037//0033-2909.112.1.155](https://doi.org/10.1037//0033-2909.112.1.155)] [Medline: [19565683](https://pubmed.ncbi.nlm.nih.gov/19565683/)]
78. Lune H, Berg BL. *Qualitative Research Methods for the Social Sciences* (9th Edition). Harlow, UK: Pearson Education Limited; 2017.

79. NVivo qualitative data analysis software. .QSR International Pty Ltd. Melbourne, Australia: QSR International; 2018. URL: <http://www.qsrinternational.com> [accessed 2021-02-10]
80. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. *Qualitative Health Research* 2005;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
81. Lewis J, Sauro J. Item Benchmarks for the System Usability Scale. *Journal of Usability Studies* 2018;13(3):158-167.
82. Niemiec RM. VIA character strengths: research and practice (the first 10 years). In: Knoop H, Delle Fave A, editors. *Well-Being and Cultures. Cross-Cultural Advancements in Positive Psychology*, vol 3. Dordrecht, The Netherlands: Springer; 2013:978-994.
83. Park N, Peterson C, Seligman MEP. Strengths of character and well-being. *Journal of Social and Clinical Psychology* 2004 Oct;23(5):603-619. [doi: [10.1521/jscp.23.5.603.50748](https://doi.org/10.1521/jscp.23.5.603.50748)]
84. Kristjansdottir OB, Børøsund E, Westeng M, Ruland C, Stenberg U, Zangi HA, et al. Mobile App to Help People With Chronic Illness Reflect on Their Strengths: Formative Evaluation and Usability Testing. *JMIR Form Res* 2020 Mar 04;4(3):e16831 [FREE Full text] [doi: [10.2196/16831](https://doi.org/10.2196/16831)] [Medline: [32130126](https://pubmed.ncbi.nlm.nih.gov/32130126/)]
85. Fredrickson BL, Joiner T. Reflections on positive emotions and upward spirals. *Perspect Psychol Sci* 2018 Mar;13(2):194-199 [FREE Full text] [doi: [10.1177/1745691617692106](https://doi.org/10.1177/1745691617692106)] [Medline: [29592643](https://pubmed.ncbi.nlm.nih.gov/29592643/)]
86. Fredrickson BL. Positive emotions broaden and build. In: *Advances in Experimental Social Psychology*, Vol 47. San Diego, CA: Academic Press; 2013:1-53.
87. Ryan RM, Deci EL. *Self-Determination Theory - Basic Psychological Needs in Motivation, Development, Wellness*. New York, NY: The Guilford Press; 2017.
88. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](https://pubmed.ncbi.nlm.nih.gov/27745683/)]
89. Hamari J, Koivisto J, Sarsa H. Does gamification work? - A literature review of empirical studies on gamification. 2014 Presented at: 2014 47th Hawaii International Conference on System Sciences; Jan 6-9, 2014; Waikoloa, HI p. 3025-3034. [doi: [10.1109/hicss.2014.377](https://doi.org/10.1109/hicss.2014.377)]
90. Cheng VWS. Recommendations for implementing gamification for mental health and qellbeing. *Front Psychol* 2020;11:586379 [FREE Full text] [doi: [10.3389/fpsyg.2020.586379](https://doi.org/10.3389/fpsyg.2020.586379)] [Medline: [33365001](https://pubmed.ncbi.nlm.nih.gov/33365001/)]
91. Amresh A, Small L. Make your garden grow: Designing a physical activity estimation improvement game. 2014 Presented at: 3rd International Conference on Serious Games and Applications for Health (SeGAH); May 14-16, 2014; Rio de Janeiro, Brazil. [doi: [10.1109/segah.2014.7067075](https://doi.org/10.1109/segah.2014.7067075)]
92. Consolvo S, Toscos T, Froehlich J, Harrison B. Activity sensing in the wild: a field trial of UbiFit Garden. 2008 Presented at: CHI '08: SIGCHI Conference on Human Factors in Computing Systems; Apr 5-10, 2008; Florence, Italy p. 1797-1806. [doi: [10.1145/1357054.1357335](https://doi.org/10.1145/1357054.1357335)]
93. Malone T. Heuristics for designing enjoyable user interfaces: Lessons from computer games. : ACM; 1982 Presented at: 1982 Conference on Human Factors in Computing Systems; March 15-17, 1982; Gaithersburg, MD p. 63-68. [doi: [10.1145/800049.801756](https://doi.org/10.1145/800049.801756)]
94. Cheng VWS, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in Apps and Technologies for Improving Mental Health and Well-Being: Systematic Review. *JMIR Ment Health* 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: [10.2196/13717](https://doi.org/10.2196/13717)] [Medline: [31244479](https://pubmed.ncbi.nlm.nih.gov/31244479/)]
95. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000 Jan;55(1):68-78. [doi: [10.1037//0003-066x.55.1.68](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
96. Sieverink F, Kelders SM, Braakman-Jansen LMA, van Gemert-Pijnen JEW. The added value of log file analyses of the use of a personal health record for patients with type 2 diabetes mellitus: preliminary results. *J Diabetes Sci Technol* 2014 Mar 04;8(2):247-255 [FREE Full text] [doi: [10.1177/1932296814525696](https://doi.org/10.1177/1932296814525696)] [Medline: [24876574](https://pubmed.ncbi.nlm.nih.gov/24876574/)]
97. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;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/)]
98. Adu MD, Malabu UH, Malau-Aduli AE, Drovandi A, Malau-Aduli BS. User retention and engagement with a mobile app intervention to support self-management in Australians with type 1 or type 2 diabetes (my care hub): mixed methods study. *JMIR Mhealth Uhealth* 2020 Jun 11;8(6):e17802 [FREE Full text] [doi: [10.2196/17802](https://doi.org/10.2196/17802)] [Medline: [32525491](https://pubmed.ncbi.nlm.nih.gov/32525491/)]
99. Cho J. The impact of post-adoption beliefs on the continued use of health apps. *Int J Med Inform* 2016 Mar;87:75-83. [doi: [10.1016/j.ijmedinf.2015.12.016](https://doi.org/10.1016/j.ijmedinf.2015.12.016)] [Medline: [26806714](https://pubmed.ncbi.nlm.nih.gov/26806714/)]
100. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. *Human-Computer Interaction* 2015 May 15;30(3-4):294-335. [doi: [10.1080/07370024.2014.993471](https://doi.org/10.1080/07370024.2014.993471)]
101. van Beugen S, Ferwerda M, Hovee D, Rovers MM, Spillekom-van Koulil S, van Middendorp H, et al. Internet-based cognitive behavioral therapy for patients with chronic somatic conditions: a meta-analytic review. *J Med Internet Res* 2014 Mar 27;16(3):e88 [FREE Full text] [doi: [10.2196/jmir.2777](https://doi.org/10.2196/jmir.2777)] [Medline: [24675372](https://pubmed.ncbi.nlm.nih.gov/24675372/)]
102. Clough BA, Casey LM. The smart therapist: a look to the future of smartphones and mHealth technologies in psychotherapy. *Professional Psychology: Research and Practice* 2015 Jun;46(3):147-153. [doi: [10.1037/pro0000011](https://doi.org/10.1037/pro0000011)]

103. Shaver LG, Khawer A, Yi Y, Aubrey-Bassler K, Etchegary H, Roebbothan B, et al. Using Facebook advertising to recruit representative samples: feasibility assessment of a cross-sectional survey. *J Med Internet Res* 2019 Aug 19;21(8):e14021 [FREE Full text] [doi: [10.2196/14021](https://doi.org/10.2196/14021)] [Medline: [31429409](https://pubmed.ncbi.nlm.nih.gov/31429409/)]
104. Marciniak MA, Shanahan L, Rohde J, Schulz A, Wackerhagen C, Kobylńska 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/)]
105. Bratteteig T. Norsk eHelse Barometer 2019. DIPS ASA. 2019. URL: https://www.dips.com/sites/default/files/norsk_ehelsebarometer_2019_-_webversjon.pdf [accessed 2021-05-10]
106. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2021 Jan 11;9(1):e21563 [FREE Full text] [doi: [10.2196/21563](https://doi.org/10.2196/21563)] [Medline: [33427672](https://pubmed.ncbi.nlm.nih.gov/33427672/)]

Abbreviations

mHealth: mobile health

SUS: System Usability Scale

Edited by G Eysenbach; submitted 20.05.21; peer-reviewed by A Terrill, H Mehdizadeh, M MacPherson, S Warsinsky; comments to author 11.06.21; revised version received 03.08.21; accepted 03.10.21; published 17.11.21.

Please cite as:

Jessen S, Mirkovic J, Halvorsen Brendmo E, Solberg Nes L

Evaluating a Strengths-Based mHealth Tool (MyStrengths): Explorative Feasibility Trial

JMIR Form Res 2021;5(11):e30572

URL: <https://formative.jmir.org/2021/11/e30572>

doi: [10.2196/30572](https://doi.org/10.2196/30572)

PMID: [34787580](https://pubmed.ncbi.nlm.nih.gov/34787580/)

©Stian Jessen, Jelena Mirkovic, Elanor Halvorsen Brendmo, Lise Solberg Nes. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.11.2021. 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 Technology-Based Pregnancy Health and Wellness Intervention (Two Happy Hearts): Case Study

Tamara Jimah¹, BSc, MA, PhD; Holly Borg¹, BA, BSc, MSN, RN; Priscilla Kehoe¹, PhD; Pamela Pimentel¹, BSc, RN; Arlene Turner², MSc; Sina Labbaf³, BSc; Milad Asgari Mehrabadi⁴, BSc, MSc; Amir M. Rahmani^{1,3,5}, MBA, PhD; Nikil Dutt^{3,4}, BE, PhD; Yuqing Guo^{1,4}, BSN, MSN, PhD, RN

¹Sue & Bill Gross School of Nursing, University of California, Irvine, Irvine, CA, United States

²First 5 Orange County Children & Families Commission, Santa Ana, CA, United States

³Department of Computer Science, University of California, Irvine, Irvine, CA, United States

⁴Department of Electrical Engineering and Computer Science, University of California, Irvine, Irvine, CA, United States

⁵Institute for Future Health, University of California, Irvine, Irvine, CA, United States

Corresponding Author:

Tamara Jimah, BSc, MA, PhD

Sue & Bill Gross School of Nursing

University of California, Irvine

299D Berk Hall

Irvine, CA, 92697

United States

Phone: 1 949 824 9057

Email: tjimah@hs.uci.edu

Abstract

Background: The physical and emotional well-being of women is critical for healthy pregnancy and birth outcomes. The Two Happy Hearts intervention is a personalized mind-body program coached by community health workers that includes monitoring and reflecting on personal health, as well as practicing stress management strategies such as mindful breathing and movement.

Objective: The aims of this study are to (1) test the daily use of a wearable device to objectively measure physical and emotional well-being along with subjective assessments during pregnancy, and (2) explore the user's engagement with the Two Happy Hearts intervention prototype, as well as understand their experiences with various intervention components.

Methods: A case study with a mixed design was used. We recruited a 29-year-old woman at 33 weeks of gestation with a singleton pregnancy. She had no medical complications or physical restrictions, and she was enrolled in the Medi-Cal public health insurance plan. The participant engaged in the Two Happy Hearts intervention prototype from her third trimester until delivery. The *Oura* smart ring was used to continuously monitor objective physical and emotional states, such as resting heart rate, resting heart rate variability, sleep, and physical activity. In addition, the participant self-reported her physical and emotional health using the Two Happy Hearts mobile app-based 24-hour recall surveys (sleep quality and level of physical activity) and ecological momentary assessment (positive and negative emotions), as well as the Perceived Stress Scale, Center for Epidemiologic Studies Depression Scale, and State-Trait Anxiety Inventory. Engagement with the Two Happy Hearts intervention was recorded via both the smart ring and phone app, and user experiences were collected via Research Electronic Data Capture satisfaction surveys. Objective data from the *Oura* ring and subjective data on physical and emotional health were described. Regression plots and Pearson correlations between the objective and subjective data were presented, and content analysis was performed for the qualitative data.

Results: Decreased resting heart rate was significantly correlated with increased heart rate variability ($r=-0.92$, $P<.001$). We found significant associations between self-reported responses and *Oura* ring measures: (1) positive emotions and heart rate variability ($r=0.54$, $P<.001$), (2) sleep quality and sleep score ($r=0.52$, $P<.001$), and (3) physical activity and step count ($r=0.77$, $P<.001$). In addition, deep sleep appeared to increase as light and rapid eye movement sleep decreased. The psychological measures of stress, depression, and anxiety appeared to decrease from baseline to post intervention. Furthermore, the participant had a high completion rate of the components of the Two Happy Hearts intervention prototype and shared several positive experiences, such as an increased self-efficacy and a normal delivery.

Conclusions: The Two Happy Hearts intervention prototype shows promise for potential use by underserved pregnant women.

KEYWORDS

ecological momentary assessment; heart rate; mHealth; physical activity; pregnancy; sleep; wearable electronic device

Introduction

The pregnancy period is characterized by physiological and psychological changes, including those affecting cardiovascular response, hormonal balance, and sleep quality [1,2]. For instance, marked cardiovascular and autonomic nervous system adaptations result in various changes, such as increased heart rate [3]. Moreover, various studies have reported a decline in sleep quality, including shorter sleep duration and increased awakenings, during pregnancy as compared to the prepregnancy period, particularly among women in the third trimester [4-6]. In addition, physical activity levels are reduced during pregnancy, and these changes are more pronounced in the third trimester [7]. All these physiological changes together may make pregnancy a stressful period, and consequently impact the emotional well-being of pregnant women. Approximately 20% of pregnant women in the United States experience prenatal distress (ie, stress, depression, or anxiety) [8], with as high as 35% experiencing elevated distress during the COVID-19 pandemic [9]. It is important to recognize that women of low socioeconomic status are at a high risk of prenatal stress [10]. Hence, effective strategies are required to promote physical and emotional well-being, healthy pregnancies, and improved birth outcomes in these vulnerable populations.

Growing research has shown that physical activity and/or mindful breathing promote positive maternal health [11-14]. However, there are mixed results concerning regular physical activity and its benefits for improving sleep and emotional well-being. A meta-analysis reported that regular exercise during pregnancy reduced the odds of self-reported prenatal depression by 67% (but not anxiety levels), with a greater reduction observed among women receiving supervision during exercise [11]. Moreover, an integrated mind-body approach such as yoga has demonstrated a decrease in perceived stress, depression, and anxiety [15]. An objective indicator of emotional state as measured by heart rate variability (HRV) is the result of cardiovascular and neurological automatic regulation [16]. Previous meta-analyses provide inconsistent evidence for the physiological impact of physical activity on HRV and sleep among pregnant women [17]. Sleep comprises an important predictor of physical and mental health [18]. Recent studies show that regular physical activity improves self-reported sleep during pregnancy, but with varying effect sizes [17,19]. Although evidence suggests a decrease in perceived prenatal distress after completion of mindful breathing practices [12,13,20], little is known about the impact of mindful breathing on sleep during pregnancy.

Existing literature suggests that mind-body interventions play an essential role in health promotion through enhanced adaptive responses during pregnancy [15]. Furthermore, coaching and social support may enhance the impact of this integrated approach to achieve emotional well-being. In line with these findings, we created the Two Happy Hearts (THH) intervention

based on the self-management theory that emphasizes an individual's active engagement in self-monitoring and reflection and, subsequently, taking steps to develop healthy behaviors [21,22]. The THH intervention is a personalized mind-body program coached by community health workers (CHWs); it includes monitoring and reflecting on personal health, as well as practicing stress management strategies, such as mindful breathing and movement. CHWs play a critical role in promoting a healthy lifestyle for pregnant women by providing compassionate listening, connecting women to community resources, and setting personalized goals and action plans tailored to their capacity and need, while supporting them to overcome potential barriers [23]. By building self-efficacy through developing and sustaining healthy habits, THH empowers women to gain self-management skills to proactively cope with stress, anxiety, and depression during pregnancy.

Self-management can be enhanced with the adoption of wearable internet of things (WIoT) technology [24]. Wearable sensors have been used to objectively assess maternal health and well-being during pregnancy given their validity and ubiquitous health monitoring [25-27]. In our study, we used the *Oura* ring, a smart ring validated to detect sleep characteristics and quality [28]. Previous research has confirmed the validity of the ring in measuring resting heart rate and HRV [29], as well as a reliable method of quantifying the level of physical activity via step count [30]. We used the *Oura* ring to collect objective data to complement self-reported responses to help pregnant women monitor their physical and emotional well-being, and to track the completion rate of the THH mindful breathing and movement components. The overarching research question was: What is the participant's usage and her perspectives regarding the THH intervention prototype? The specific aims of the study were to (1) test the daily use of a wearable device to objectively measure physical and emotional well-being along with subjective assessments during pregnancy, and (2) explore the user's engagement with the THH intervention prototype, as well as understand their experiences with the prototype components.

Methods

Study Design

This case study design involved a single participant receiving the THH intervention prototype. Both quantitative and qualitative data were collected using a wearable device and by administering closed- and open-ended survey questionnaires.

THH Intervention Prototype

We developed the THH intervention prototype according to the American College of Obstetricians and Gynecologists (ACOG) exercise guidelines and updated evidence on mindful breathing [12,31,32]. Mindful breathing involved paced deep breathing (ie, 3 minutes per practice, twice a day). Mindful movement included brisk walking (ie, 30 minutes on most days), and strength movement (ie, squat, plie squat, wall push-up, one-arm

row, bicep curl, overhead press, and low row) with 12 to 15 repetitions per movement (3 times per week). Walking and strength movement started and ended with 2 minutes of paced deep breathing. All study visits were conducted virtually by a trained CHW. The CHW assessed the participant's physical and mental health history, helped set personalized goals, created an action plan suitable for the participant's health condition, fitness level and schedule, as well as provided continuous support via text messages and/or brief phone conversations.

A unique feature of the THH intervention prototype was the leveraging of technology to strengthen emotional and social support provided to the participant. The *Oura* ring was used to enhance the participant's awareness of self-monitoring of vital

health indicators, such as resting heart rate, resting HRV, sleep, and physical activity. Moreover, our research team designed the THH mobile phone app to enhance the participant's experience with the intervention. The THH app was developed using ZotCare, a health cybernetic service platform that makes it possible to build a cybernetic system for data collection, labeling, and intervention. Specifically, the app-based surveys helped with self-reflection on physical and emotional well-being. In addition, video demonstrations and guidelines were incorporated into the app to increase the user's access to health-related education material on mindful breathing and movement. Figure 1 shows the different components of the THH intervention prototype.

Figure 1. Components of the Two Happy Hearts intervention prototype.



Study Setting and Participant Recruitment

We recruited 1 low-income pregnant woman from MOMS Orange County (MOMS OC), a community-based nonprofit organization that supports underserved perinatal women in Orange County, California. The participant received the study flyer from MOMS OC and contacted the THH research team. She was subsequently screened by the study coordinator for the following eligibility criteria: adult aged 18-40 years, having a singleton pregnancy without medical complications restricting physical activity, not engaged in a coached physical activity and/or mindfulness program, and having a smartphone. The participant consented to engage in the THH intervention prototype facilitated by a CHW virtually, as required per COVID-19 pandemic regulations.

Data Collection

Quantitative Data

Objective data were obtained from the *Oura* ring. Subjective data (ie, responses to closed-ended questions) were collected via the app-based surveys and the Research Electronic Data Capture (REDCap), a secure web-based data collection survey tool [33]. The app-based surveys included sleep recall, physical

activity recall, and ecological momentary assessment (EMA). The REDCap collected pre-post intervention surveys at 33 weeks of gestation, before the start of the intervention, and at 1 week post partum.

Qualitative Data

Qualitative data were obtained from open-ended questions about the participant's user experience with the THH intervention. The CHW recorded the participant's responses on REDCap during the interview. Additional feedback was received from the participant via email.

Measures

Objective Measures of Physical and Emotional Well-being

For the objective data, we used within-subject repeated measures to monitor physical and emotional well-being. Cardiovascular response (ie, average resting heart rate and HRV), sleep, and physical activity data were collected daily using the *Oura* ring [34]. Resting HRV was calculated in terms of the root mean square of successive difference (RMSSD) of interbeat intervals. Two attributes of sleep were selected for our analysis: sleep stages and sleep score. Sleep stages include deep, rapid eye

movement (REM), and light sleep. Sleep score ranges from 0 to 100 and denotes the quality of sleep [34]. The following are suggested sleep score ranges to determine sleep quality: <70 (pay attention to sleep); ≥70 to 84 (good); and ≥85 (optimal) [34]. Step-count data from the ring was used as an indicator of physical activity level. The frequency and duration of the THH mindful breathing and strength movement practice were recorded with the THH mobile app.

Subjective Measures of Physical and Emotional Well-being

The demographic survey was administered via REDCap to understand the participant's sociodemographic background, physical and mental health history, and current pregnancy-related conditions. Three validated psychological measures that were previously tested with acceptable reliability for studies involving pregnant women were used to assess stress, depression, and anxiety [35-37].

The Perceived Stress Scale (PSS) [38,39] measures the frequency of stress experienced over the past month on a 5-point Likert scale, ranging from 0 (Never) to 4 (Very often). An example of the PSS is "In the last month, how often have you been upset because of something that happened unexpectedly?" The PSS comprised 10 items with scores ranging from a minimum of 0 to a maximum of 40 points.

The Center for Epidemiologic Studies Depression Scale (CES-D) [40] was used to measure depression symptoms on a 4-point Likert scale, ranging from 0 (Rarely or none of the time: less than 1 day) to 3 (Most or all of the time: 5-7 days). An example question of the CES-D is "In the past week, I was bothered by things that usually do not bother me." The CES-D included 20 items, with scores ranging from 0 to 60 points.

The State Trait Anxiety Inventory (STAI) [41] was adopted to measure the extent of anxiety-related symptoms or emotions on a 4-point Likert scale, ranging from 1 (Not at all) to 4 (Very much). STAI includes 20 items with scores ranging from a minimum of 20 to a maximum of 80 points. An example question of the STAI is "I feel calm at this exact moment."

A 24-hour recall survey was administered daily in the morning using the THH mobile app to enhance pregnant women's awareness of personal health. The survey comprised 2 questions: (1) "How would you rate your quality of sleep last night?" with response options 0 (Very poor), 1 (Poor), 2 (Fair), 3 (Good), and 4 (Very good); and (2) "How physically active were you in the past 24 hours?" with response options 0 (Not at all), 1 (A little), 2 (Somewhat), 3 (Quite a bit), and 4 (A lot).

The EMA [42] was administered daily around noontime using the THH mobile app. Its purpose was to enhance the participant's awareness of her emotional states and to empower her to manage negative emotions by engaging in mindful breathing and physical activity. The EMA included 7 items capturing both positive and negative emotions. Each item score ranged from 0 (Not at all) to 4 (A lot). The questions related to positive emotions assessed the participant's sense of health and safety, with scores ranging from 0 to 8. The questions related to negative emotions assessed the participant's feelings of

depression, stress, worry, anger, and loneliness, with scores ranging from 0 to 20.

User Experience

User experience was measured using the REDCap closed- and open-ended survey items to assess participant's experiences with the smart ring, app-based survey, as well as the THH mindful breathing and movement components. An example of a closed-ended survey item was "Practicing THH Mindful Breathing helped me manage stress in my daily life" Response choices were 0 (Completely disagree), 1 (Disagree), 2 (Neither agree nor disagree), 3 (Agree), and 4 (Completely agree). Open-ended questions sought to understand the user's feedback regarding the THH intervention prototype, for example, "What was your favorite feature? Are there any changes you would suggest for improvement?"

Analyses

Quantitative Analysis

Statistical analyses were conducted using Stata software (version 15.1; StataCorp LLC). Normality was determined for the objective and subjective measures using histograms and skewness tests. Regression plots with 95% CIs were used to illustrate objective measures (eg, resting heart rate, resting HRV, sleep score, stages of sleep, and steps) and subjective measures (eg., positive/negative emotions, sleep quality, physical activity level) from 33 to 40 weeks of gestation (Figures 2-12). Given that we sought to understand the correlations between objective and subjective data, subjective data were recoded numerically [43]. Subsequently, Pearson correlation tests were conducted to examine the associations between objective and subjective measures of physical and emotional well-being. The sum scores of the pre-post THH psychological measures were calculated, with reverse coding applied where necessary. A higher PSS, CES-D, and STAI score indicated greater stress, depression, and anxiety level, respectively. The frequency, duration, and completion rate of each THH intervention prototype component were also reported.

Qualitative Analysis

For the qualitative analysis, we reviewed the responses to the open-ended questions using content analysis, that is, a research method providing perceived meanings and perspectives [44,45]. Coding and organization of the qualitative data involved 3 steps: (1) review of responses to understand the participant's perspectives; (2) deriving codes that capture the participant's experiences; and (3) summarizing codes into themes [44].

Ethical Considerations

Prior to recruitment and data collection, study approval was obtained from the institutional review board at the University of California, Irvine. The study participant was screened for eligibility and provided with a study information sheet outlining her rights and participation requirements. Data collection began after receiving the participant's consent.

Results

Participant Characteristics

The participant was in her third trimester at the time of enrollment; she was a first-time mother without any medical and pregnancy conditions restricting exercise. She self-identified as Hispanic and Caucasian, preferred to speak English at home, had completed some college education, and worked in an administrative position. At the time of the study, she was enrolled in the Medi-Cal public health insurance plan, was divorced from the father of the baby and living with a male partner. Prior to her pregnancy, she exercised 5 to 6 times a week, but had reduced that to 3 to 4 times during early pregnancy. Factors that motivated her most to exercise included her desire for a healthy baby, healthy pregnancy, and improved mental health. However, she was unable to maintain her regular exercise routine due to the stay-at-home order pertaining to COVID-19 restrictions until she enrolled into the THH intervention. The total duration of her participation in the THH intervention was 8 weeks (ie, 33-40 weeks of gestation). Data were collected from April 18, 2020, to June 12, 2020.

Quantitative Results

Objective and Subjective Measures of Physical and Emotional Well-being

The first aim of the study involved testing the objective and subjective measures of well-being from the third trimester of pregnancy until delivery. Figures 2 and 3 illustrate the cardiovascular response patterns over the course of the final trimester of pregnancy. Decreased resting heart rate was significantly correlated with increased HRV ($r=-0.92$, $P<.001$). Figures 4 and 5 show the positive and negative emotions over time, as reported by the participant. Increased positive emotions were positively correlated with HRV, an objective indicator of emotional states ($r=0.54$, $P<.001$). The correlation between negative emotions and HRV was not significant ($r=0.18$, $P=.22$). Self-reported quality of sleep (Figure 6) was positively correlated with the sleep score generated from the *Oura* ring ($r=0.52$, $P<.001$; Figure 7). In terms of the stages of sleep, deep sleep (Figure 8) appeared to increase, whereas light sleep and REM sleep appeared to decrease (Figures 9 and 10). There was a gradual increase in the level of self-reported physical activity (Figure 11), which was positively correlated with steps (Figure 12) from the smart ring ($r=0.77$, $P<.001$).

Figure 2. Average resting heart rate. HR: heart rate.

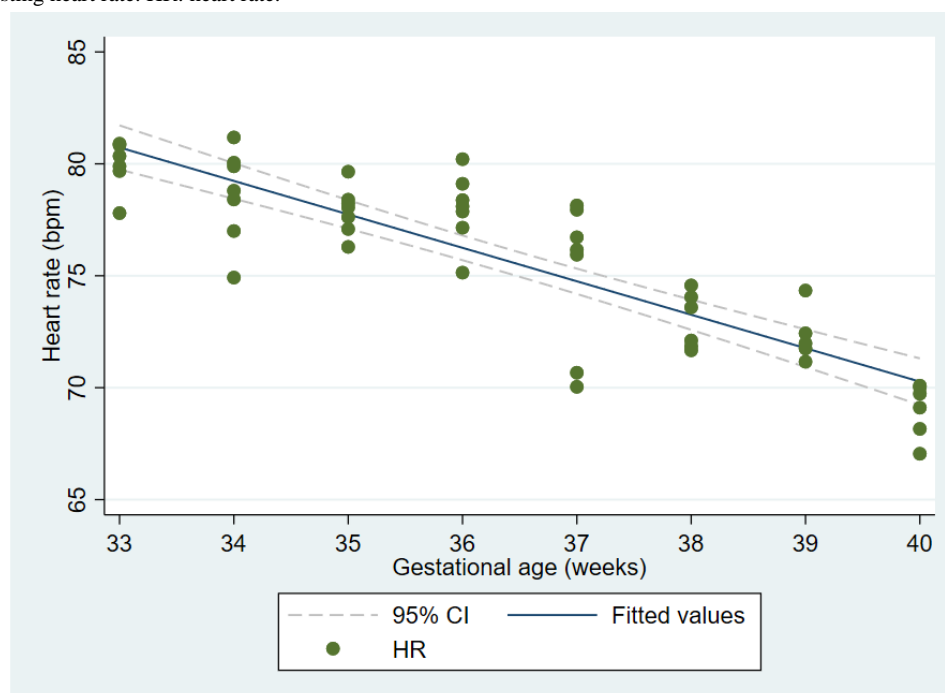


Figure 3. Average resting heart rate variability. RMSSD: root mean square of successive differences.

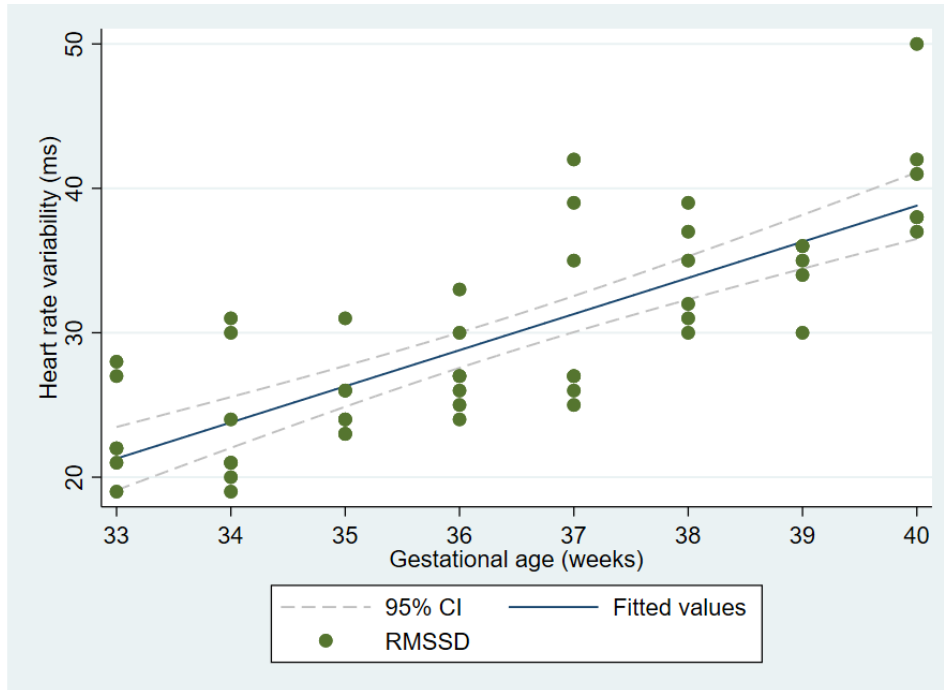


Figure 4. Self-reported positive emotions. EMA: ecological momentary assessment.

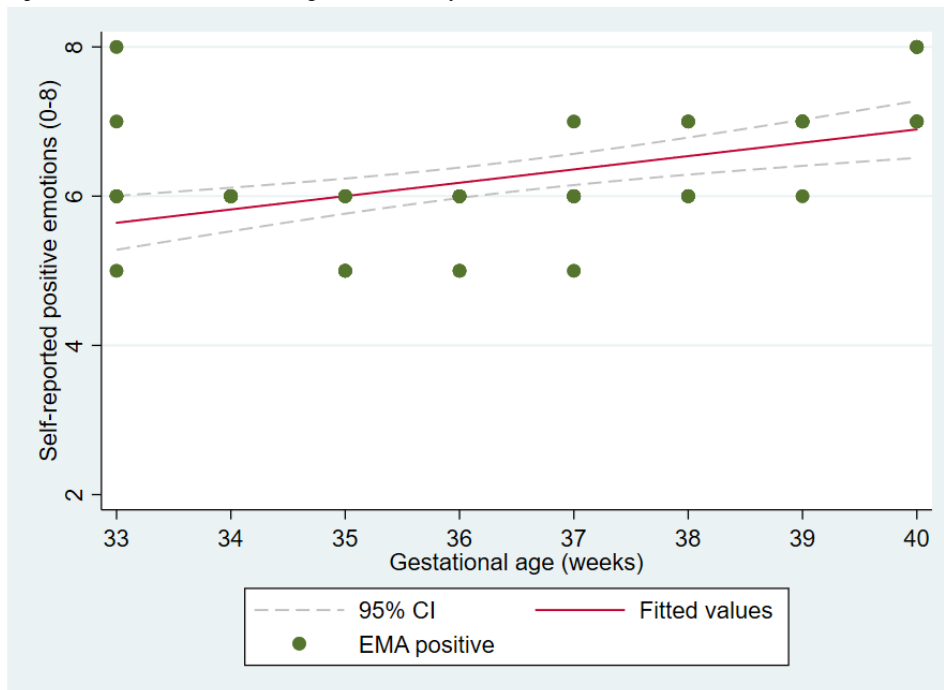


Figure 5. Self-reported negative emotions. EMA: ecological momentary assessment.

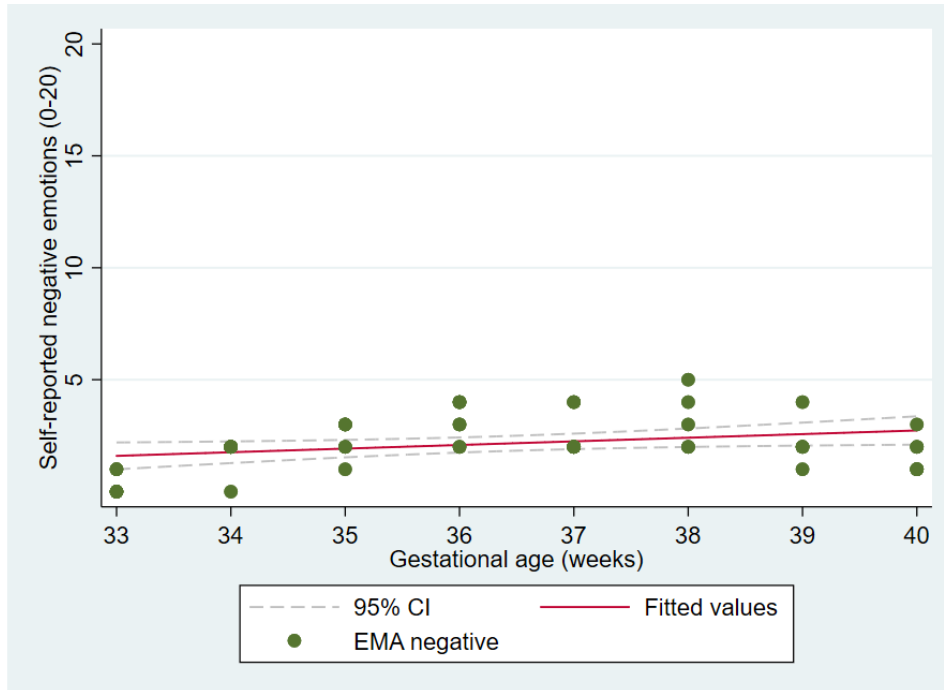


Figure 6. Self-reported quality of sleep.

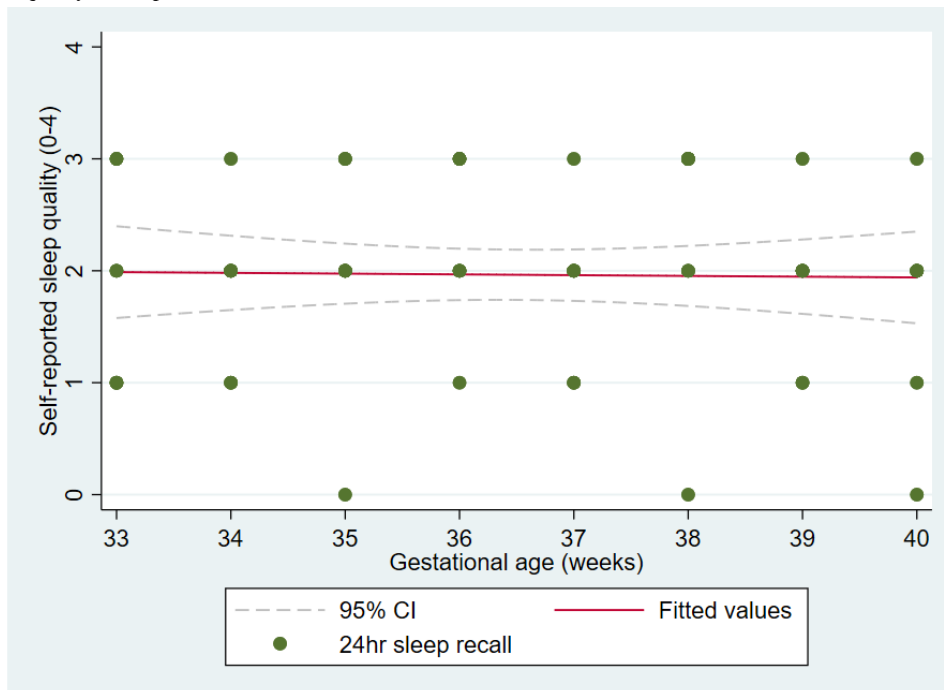


Figure 7. Sleep score as measured by the Oura ring.

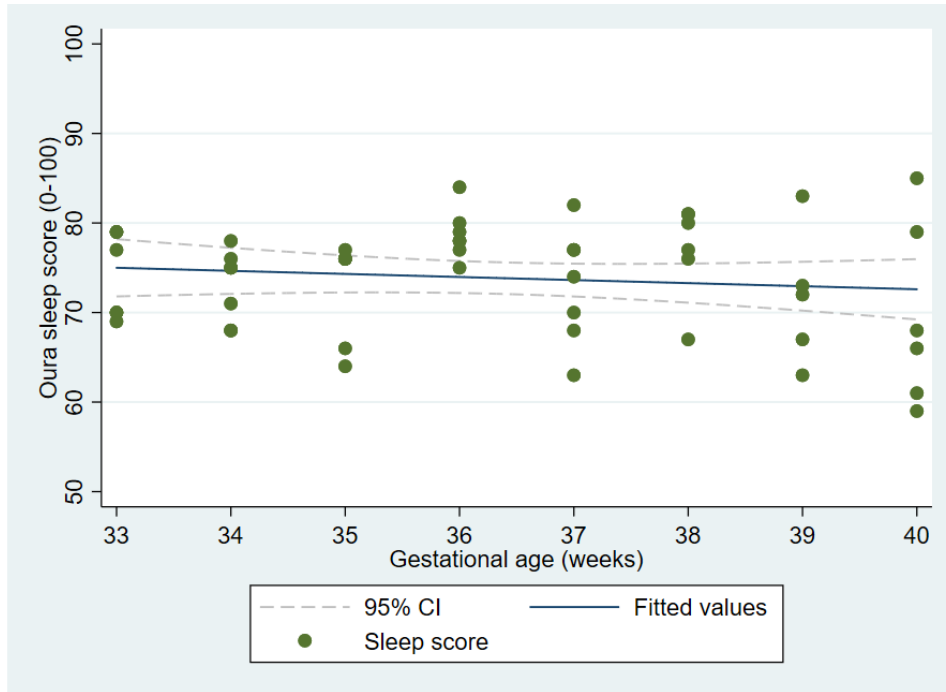


Figure 8. Stage of sleep (deep sleep).

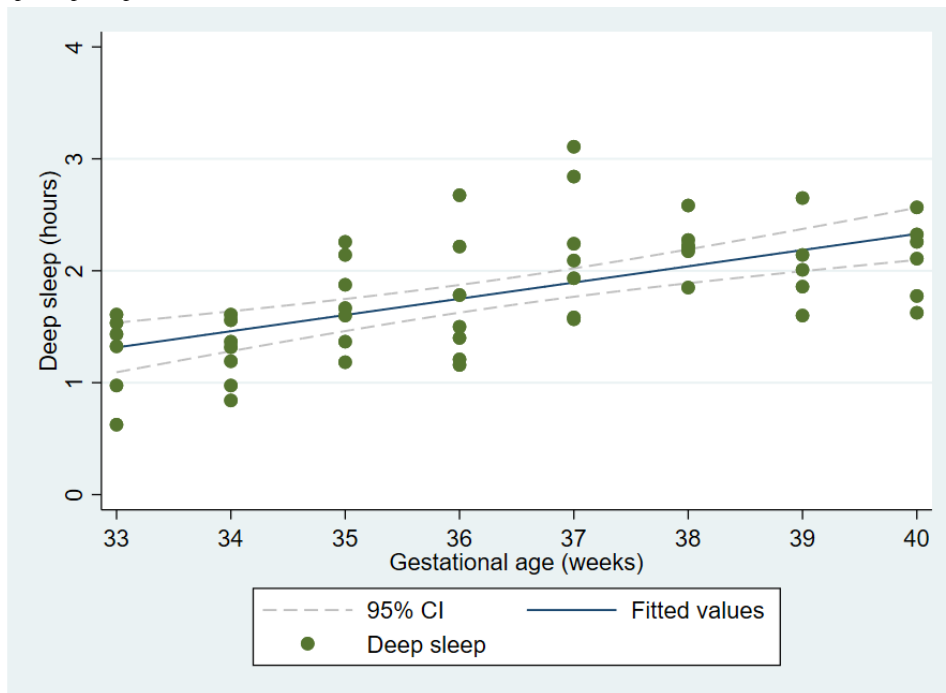


Figure 9. Stage of sleep (light sleep).

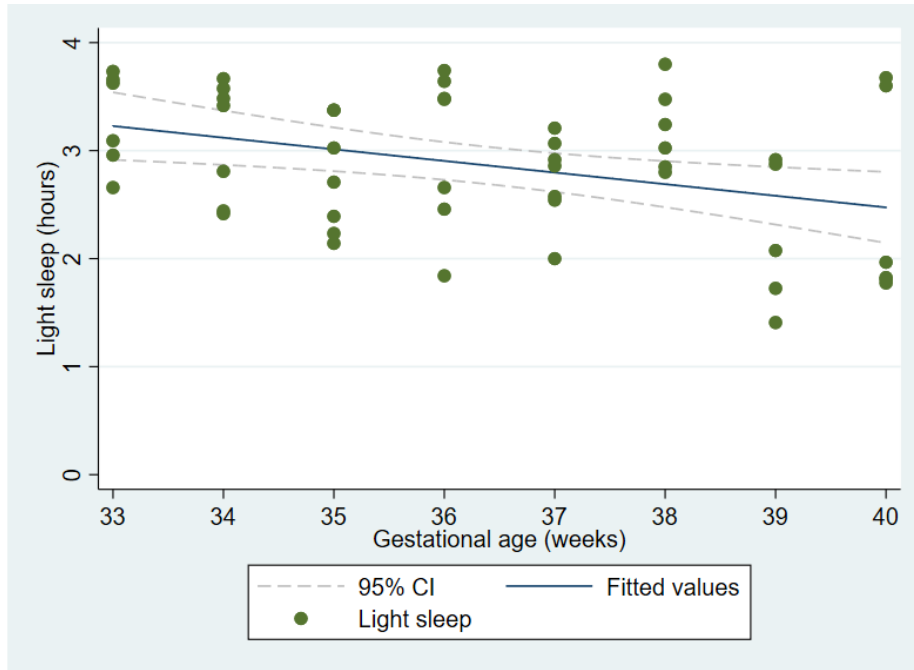


Figure 10. Stage of sleep (rapid eye movement, or REM sleep).

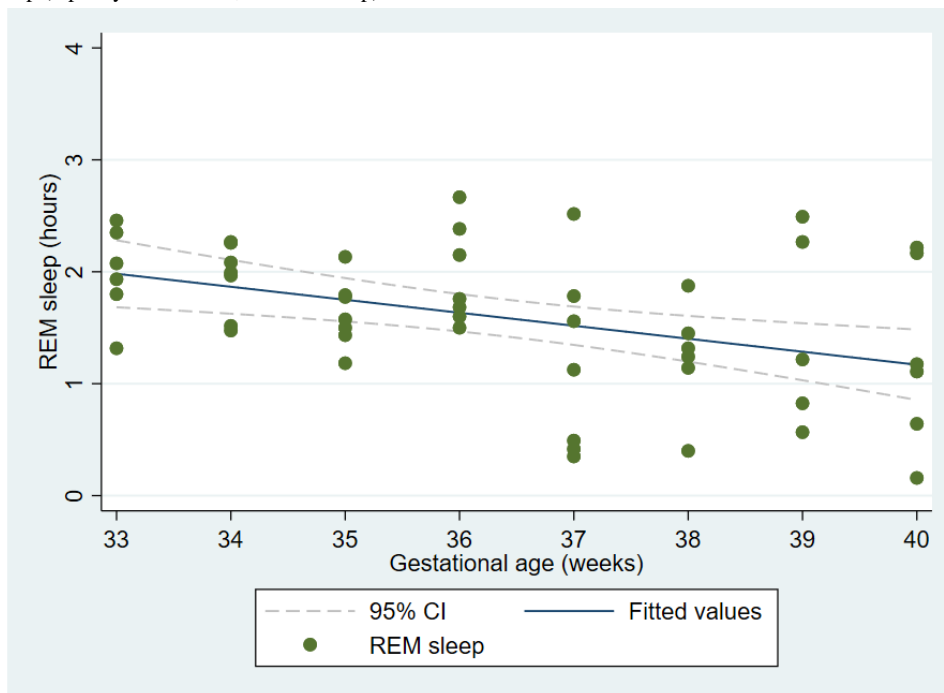


Figure 11. Self-reported level of physical activity.

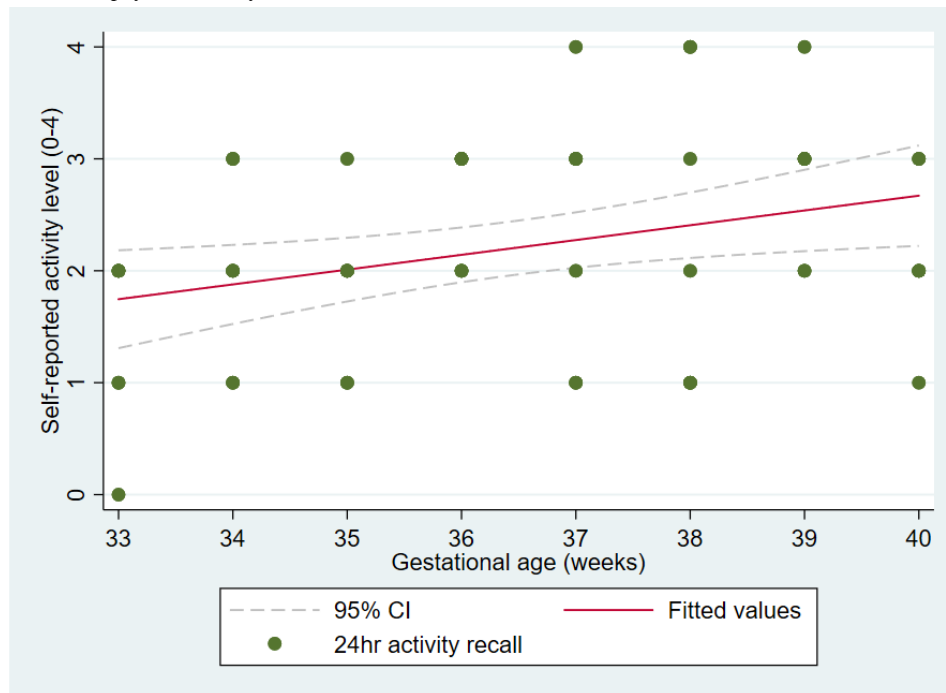
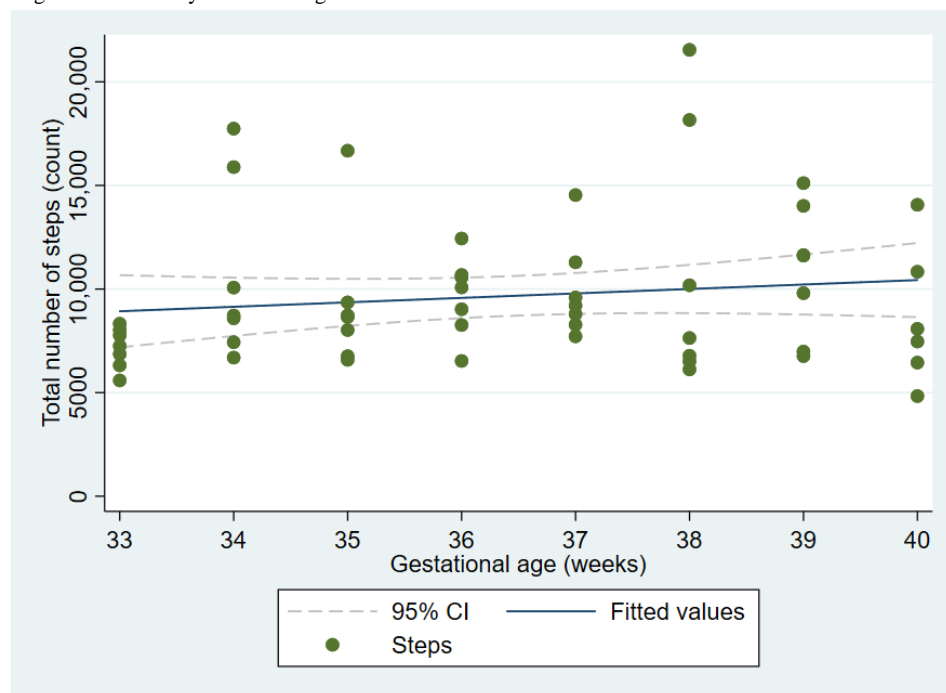


Figure 12. Step recordings as measured by the Oura ring.



Experience With the THH Intervention Prototype

The second aim of the study sought to understand the participant’s engagement with the THH intervention prototype and her overall user experience. Table 1 shows the participant’s

engagement (ie, frequency and duration) and high completion rate (94.6%-100%) of the THH intervention prototype components, specifically the smart ring, app-based surveys, and mind-body movement..

Table 1. Engagement and completion of the Two Happy Hearts (THH) intervention prototype.

THH intervention components and their description	Frequency or duration (% completion rate)
Smart Ring	
Resting heart rate, resting heart rate variability, sleep, step output	Daily for 56 days (100) ^a
App-based surveys	
24-hour recall of general health	Once a day for 56 days (100) ^a
Ecological momentary assessment	Once a day for 53 days (94.6) ^{a,b}
Mind-body movement	
Walking (step count)	Average of 9666 steps a day for 55 days (98.2) ^{a,c}
Mindful breathing	Twice a day (3 minutes per practice) for 2 weeks (100) ^a
Strength movement	Three times per week for 5 weeks (100) ^a

^aTotal days in the THH intervention prototype=56 days. Components of the intervention were brought in at different times.

^bThree days of missing data.

^cOutlier excluded.

Table 2 shows the observed changes before and after completing the THH intervention prototype. All psychological measures, including PSS, CES-D, and STAI scores, appeared to decrease from baseline to postintervention. The participant rated her satisfaction with the THH intervention an average score of 3.5 (SD 0.9) out of a total of 4 points. She particularly enjoyed

using the smart ring, completing the app-based surveys, and mindful strength movement, indicating that the ring was comfortable and easy to use, completing the app-based questions did not interfere with her daily life, and that the strength movement helped her to have a normal delivery.

Table 2. Psychological measures before and after the Two Happy Hearts (THH) intervention prototype.

Psychological measures	Baseline score (at 33 weeks of gestation)	Postpartum reduction of score (1 week after delivery)
Perceived Stress Scale (max score = 40)	11	6
Center for Epidemiologic Studies Depression Scale ^a (max score = 60)	17	7
State Trait Anxiety Inventory ^b (max score = 80)	24	20

^aA score of ≥ 16 indicates depression [46].

^bA score of ≥ 45 indicates anxiety [36].

Qualitative Results

Regarding user experience, the study participant described the intervention as positive and beneficial. Responses to the open-ended surveys were summarized under the 4 main themes described below.

Social Support

The participant noted the care and support received from her THH health coach, saying the following:

...feels like talking with a sister or good friend. I knew I was working with someone who genuinely cared, and had my best interests in mind. We worked together for several weeks to better the programming and fine-tune the comfort of the plan I chose.

Health Tracking

The THH mobile app was easy to navigate, and the participant enjoyed using the app to practice her exercise routines. The participant also appreciated the ability to self-monitor her sleep and activity level on a daily basis using the smart ring, noting

that this function allowed her to effectively make more informed health choices. For example, during one of the sessions with her CHW, she shared that she wanted to be careful “*not birth a ten-pound baby and nine-pound placenta’ and be out of her weight range after delivery.*” Moreover, she valued the helpful information provided by the smart ring, such as suggestions on improving her sleep habits.

Well-being and Birth

Our participant stated that she enjoyed practicing the THH strength movement and was confident in her ability to safely engage in the recommended exercise routines. She particularly liked that the mind-body movement helped her to feel more relaxed and sleep better. During labor, she performed 60 squats while having contractions, which she believed helped her to dilate and deliver *quickly*. During the exit interview with her CHW, she stated the following:

I couldn’t have done any of this without the program and probably would have opted for epidural had I not stayed so active...I will continue to implement some of the moves shown to me during my

workout...I've caught myself taking a moment to breathe and can hear the clip in my head directing me and it has reminded me to pause every once in a while during my crazy busy days.

Recommendations

The participant provided some suggestions for improving the study. For example, she noted experiencing some struggle with mindful breathing given that the THH app had only audio cues and no visual cues or animation. She would also have preferred additional calendar notifications for mindful breathing and movement. She suggested that integrating the THH platform with her prenatal health data would enable her to share this information with her health care provider during prenatal visits.

Discussion

Principal Results

Objective and Subjective Measures of Physical and Emotional Well-being

We were able to use the *Oura* ring to measure objective indicators of physical and emotional well-being from mid-third trimester until delivery. There was a decrease in resting heart rate and an increase in HRV. Deep sleep appeared to increase as light and REM sleep decreased. In addition, we found significant positive associations between self-reported well-being and *Oura* ring measures- higher positive emotions with greater HRV, better sleep quality with a higher sleep score, as well as increased physical activity with a higher step count.

Experience With the THH Intervention Prototype

Our participant effectively engaged with the THH intervention as observed by her consistency in wearing the smart ring, completing the app-based surveys, and practicing the mindful breathing and movement routines. Her feedback indicated a positive and rewarding experience with the THH intervention. Specifically, she found it useful to complete the app-based surveys to better understand her physical and emotional well-being, felt confident to practice the strength movement routines safely and attributed her uncomplicated and *quick* labor and delivery to the intervention.

Comparison With Prior Work

Objective and Subjective Measures of Physical and Emotional Well-being

We found that HRV had a significant positive correlation with positive emotions, but not with negative emotions. There are inconsistent findings about these associations in the existing literature. In one study, no significant association was found between fear of childbirth and HRV among pregnant women at 32 to 34 weeks of gestation [47]. However, higher self-perceived stress and lower HRV was reported among pregnant women between 12 and 30 weeks of gestation [48]. The literature also shows mixed results about HRV patterns during pregnancy, with and without intervention [49-51]. It is worth noting that various factors may influence self-reported emotions and HRV, including pregnancy progression, external circumstances such as the COVID-19 pandemic, and different

methods of measuring HRV (eg, wearable device in the home setting vs an electrocardiogram in the laboratory), as well as temporal factors taken into account during measurement (eg, continuous vs short-term interval and day vs night) [28,49,50,52,53]. Our findings may have important implications for continuous remote monitoring of objective and subjective emotional health, particularly during heightened levels of stress and anxiety over the course of pregnancy.

With regard to sleep, we observed a decline in rapid eye movement during sleep, as previously reported [5,54]. Deep sleep is an indicator of sleep quality [55]. Interestingly, our study found an increase in deep sleep and a decrease in light sleep. These results are contrary to those of prior observational research involving pregnant women who did not participate in any health intervention during the third trimester [56] and suggest potential benefits of the THH intervention prototype. Additionally, self-reported sleep quality from the app-based 24-hour recall surveys corresponded with the sleep score extracted from the smart ring. Furthermore, self-reported physical activity was consistent with the step count recorded by the ring. Overall, our results suggest that WIoT, such as the *Oura* ring, along with self-reported data from app-based surveys, may provide important objective and subjective measures of physical and emotional well-being during pregnancy.

Experience With the THH Intervention Prototype

Recent literature emphasizes the opportunity to leverage technology to promote physical and emotional well-being for pregnant women [57,58]. The high completion rate of the intervention components and positive feedback from our participant shows promise for potential use by underserved pregnant women, many of whom lack access to quality health information, physical activity instruction, stress reduction techniques, and other pregnancy-related resources. Our results align with those of a recent study that found a high adherence rate and acceptance of a mobile health-based physical activity intervention among underserved pregnant women [59]. On the other hand, it is worth noting technology-related challenges that may prevent effective engagement in exercise interventions for pregnant women. For example, time constraints, low technology literacy, and design or complexity of the user interface comprise few factors that have been observed in this study and in previous research [57,60]. Results from our case study are promising, and we anticipate refining the THH intervention to incorporate lessons learned as we support pregnant women to use self-management strategies for improved health and well-being.

Strengths and Limitations

Our study has several strengths. First is the use of both objective and subjective measures to monitor health and well-being during pregnancy. The *Oura* ring features allowed us to continuously assess important parameters of cardiovascular responses, sleep, and physical activity. In addition, the THH phone app-based surveys captured self-reported health and emotional states daily. Our results suggest the potential of the smart ring and app-based surveys to monitor physical and emotional states. The enthusiasm and commitment demonstrated by the participant's completion of the THH intervention prototype components was

especially valuable given that COVID-19 pandemic regulations made in-person engagement impossible. As seen from her feedback and experience, the intervention provided her with care and support during her pregnancy. Our participant was of low socioeconomic status, and while she was motivated to exercise, THH provided her with the opportunity to do so safely. Therefore, the THH intervention has the potential to reach women in underserved communities, who often face significant barriers to essential health information and feedback about healthy lifestyle choices during pregnancy. It is therefore critical to refine the THH intervention prototype, integrating our participant's suggestions, as well as test the study's feasibility in a large sample.

Limitations and possible areas for improvement must be noted. For example, data were available for the third trimester only, which did not allow for comparisons of health measures between trimesters. Recruiting women at earlier stages of pregnancy to examine how physical and emotional health and well-being vary across trimesters is recommended for future research.

External factors may have influenced these study results. For instance, the participant was recruited from MOMS OC where she received additional social support during her pregnancy. Also, although the participant desired to have had her final post-intervention interview prior to giving birth, she was unable to complete it until a week following delivery due to time constraints related to birth preparations. Finally, this is a case study; thus, the study results are not generalizable.

Conclusions

To our knowledge, this study is the first to continuously monitor objective and subjective health and well-being in a pregnant woman from the third trimester until delivery. The consistent patterns between the app-based self-reported data and the *Oura* ring data suggest the potential use of the smart ring to provide valuable health data for pregnant women. The enthusiasm and commitment demonstrated by the participant's completion of the THH intervention prototype was especially promising. It is, therefore, important to refine the intervention and subsequently test its feasibility among underserved pregnant women.

Acknowledgments

The following contributors are to be considered co-corresponding authors:

Tamara Jimah, BSc, MA, PhD, Email: tjimah@hs.uci.edu

Yuqing Guo, BSN, MSN, PhD, RN, Email: gyuqing@hs.uci.edu

This research was supported by the US National Science Foundation, under the Smart and Connected Communities (S&CC) (grant CNS-1831918). The authors would like to thank MOMS Orange County for help with recruitment, and First 5 Orange County Children & Families Commission for the support provided.

Authors' Contributions

YG, PK, PP, AT, AR, and ND designed and supervised the study. HB, PP, and SL collected the data. TJ, YG, and PK conceived and designed the manuscript. TJ, MA, YG, and PK conducted the data analysis. TJ, YG, and PK wrote the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Bjelica A, Cetkovic N, Trninic-Pjevic A, Mladenovic-Segedi L. The phenomenon of pregnancy - a psychological view. *Ginekol Pol* 2018;89(2):102-106 [[FREE Full text](#)] [doi: [10.5603/GPa2018.0017](https://doi.org/10.5603/GPa2018.0017)] [Medline: [29512815](https://pubmed.ncbi.nlm.nih.gov/29512815/)]
2. Matvienko-Sikar K, Dockray S. Effects of a novel positive psychological intervention on prenatal stress and well-being: A pilot randomised controlled trial. *Women Birth* 2017 Apr;30(2):e111-e118. [doi: [10.1016/j.wombi.2016.10.003](https://doi.org/10.1016/j.wombi.2016.10.003)] [Medline: [27810284](https://pubmed.ncbi.nlm.nih.gov/27810284/)]
3. Heiskanen N, Saarelainen H, Kärkkäinen H, Valtonen P, Lyyra-Laitinen T, Laitinen T, et al. Cardiovascular autonomic responses to head-up tilt in gestational hypertension and normal pregnancy. *Blood Press* 2011 Apr;20(2):84-91. [doi: [10.3109/08037051.2010.532313](https://doi.org/10.3109/08037051.2010.532313)] [Medline: [21105762](https://pubmed.ncbi.nlm.nih.gov/21105762/)]
4. Hutchison BL, Stone PR, McCowan LME, Stewart AW, Thompson JMD, Mitchell EA. A postal survey of maternal sleep in late pregnancy. *BMC Pregnancy Childbirth* 2012 Dec 10;12:144 [[FREE Full text](#)] [doi: [10.1186/1471-2393-12-144](https://doi.org/10.1186/1471-2393-12-144)] [Medline: [23228137](https://pubmed.ncbi.nlm.nih.gov/23228137/)]
5. Izci-Balsarak B, Keenan BT, Corbitt C, Staley B, Perlis M, Pien GW. Changes in sleep characteristics and breathing parameters during sleep in early and late pregnancy. *J Clin Sleep Med* 2018 Jul 15;14(7):1161-1168 [[FREE Full text](#)] [doi: [10.5664/jcsm.7216](https://doi.org/10.5664/jcsm.7216)] [Medline: [29991418](https://pubmed.ncbi.nlm.nih.gov/29991418/)]
6. Polo-Kantola P, Aukia L, Karlsson H, Karlsson L, Paavonen EJ. Sleep quality during pregnancy: associations with depressive and anxiety symptoms. *Acta Obstet Gynecol Scand* 2017 Feb;96(2):198-206 [[FREE Full text](#)] [doi: [10.1111/aogs.13056](https://doi.org/10.1111/aogs.13056)] [Medline: [27861713](https://pubmed.ncbi.nlm.nih.gov/27861713/)]

7. Downs DS, Chasan-Taber L, Evenson KR, Leiferman J, Yeo S. Physical activity and pregnancy: past and present evidence and future recommendations. *Res Q Exerc Sport* 2012 Dec;83(4):485-502 [FREE Full text] [doi: [10.1080/02701367.2012.10599138](https://doi.org/10.1080/02701367.2012.10599138)] [Medline: [23367811](https://pubmed.ncbi.nlm.nih.gov/23367811/)]
8. Obrochta CA, Chambers C, Bandoli G. Psychological distress in pregnancy and postpartum. *Women Birth* 2020 Nov;33(6):583-591. [doi: [10.1016/j.wombi.2020.01.009](https://doi.org/10.1016/j.wombi.2020.01.009)] [Medline: [32035798](https://pubmed.ncbi.nlm.nih.gov/32035798/)]
9. Liu CH, Erdei C, Mittal L. Risk factors for depression, anxiety, and PTSD symptoms in perinatal women during the COVID-19 Pandemic. *Psychiatry Res* 2021 Jan;295:113552. [doi: [10.1016/j.psychres.2020.113552](https://doi.org/10.1016/j.psychres.2020.113552)] [Medline: [33229122](https://pubmed.ncbi.nlm.nih.gov/33229122/)]
10. Lefmann T, Combs-Orme T, Orme JG. Examining the inter-correlated effects of low income, life stress, and race on birth outcomes: A representative state study. *Soc Work Health Care* 2017 Jul;56(6):450-469. [doi: [10.1080/00981389.2017.1316811](https://doi.org/10.1080/00981389.2017.1316811)] [Medline: [28486063](https://pubmed.ncbi.nlm.nih.gov/28486063/)]
11. Davenport MH, McCurdy AP, Mottola MF, Skow RJ, Meah VL, Poitras VJ, et al. Impact of prenatal exercise on both prenatal and postnatal anxiety and depressive symptoms: a systematic review and meta-analysis. *Br J Sports Med* 2018 Nov;52(21):1376-1385. [doi: [10.1136/bjsports-2018-099697](https://doi.org/10.1136/bjsports-2018-099697)] [Medline: [30337464](https://pubmed.ncbi.nlm.nih.gov/30337464/)]
12. Fiskin G, Sahin N. Effect of diaphragmatic breathing exercise on psychological parameters in gestational diabetes: A randomised controlled trial. *Eur J Integr Med* 2018 Oct;23:50-56 [FREE Full text] [doi: [10.1016/j.eujim.2018.09.006](https://doi.org/10.1016/j.eujim.2018.09.006)]
13. Chang S, Kim H, Ko Y, Bae C, An S. Effects of abdominal breathing on anxiety, blood pressure, peripheral skin temperature and saturation oxygen of pregnant women in preterm labor. *Korean J Women Health Nurs* 2009;15(1):32-42 [FREE Full text] [doi: [10.4069/kjwhn.2009.15.1.32](https://doi.org/10.4069/kjwhn.2009.15.1.32)]
14. Davenport MH, Sobierajski F, Mottola MF, Skow RJ, Meah VL, Poitras VJ, et al. Glucose responses to acute and chronic exercise during pregnancy: a systematic review and meta-analysis. *Br J Sports Med* 2018 Nov;52(21):1357-1366. [doi: [10.1136/bjsports-2018-099829](https://doi.org/10.1136/bjsports-2018-099829)] [Medline: [30337462](https://pubmed.ncbi.nlm.nih.gov/30337462/)]
15. Kwon R, Kasper K, London S, Haas DM. A systematic review: The effects of yoga on pregnancy. *Eur J Obstet Gynecol Reprod Biol* 2020 Jul;250:171-177. [doi: [10.1016/j.ejogrb.2020.03.044](https://doi.org/10.1016/j.ejogrb.2020.03.044)] [Medline: [32446148](https://pubmed.ncbi.nlm.nih.gov/32446148/)]
16. Kim H, Cheon E, Bai D, Lee YH, Koo B. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry Investig* 2018 Mar;15(3):235-245 [FREE Full text] [doi: [10.30773/pi.2017.08.17](https://doi.org/10.30773/pi.2017.08.17)] [Medline: [29486547](https://pubmed.ncbi.nlm.nih.gov/29486547/)]
17. Kredlow MA, Capozzoli MC, Hearon BA, Calkins AW, Otto MW. The effects of physical activity on sleep: a meta-analytic review. *J Behav Med* 2015 Jun;38(3):427-449. [doi: [10.1007/s10865-015-9617-6](https://doi.org/10.1007/s10865-015-9617-6)] [Medline: [25596964](https://pubmed.ncbi.nlm.nih.gov/25596964/)]
18. Christian LM, Carroll JE, Teti DM, Hall MH. Maternal sleep in pregnancy and postpartum part I: mental, physical, and interpersonal consequences. *Curr Psychiatry Rep* 2019 Mar 02;21(3):20. [doi: [10.1007/s11920-019-0999-y](https://doi.org/10.1007/s11920-019-0999-y)] [Medline: [30826881](https://pubmed.ncbi.nlm.nih.gov/30826881/)]
19. Yang S, Lan S, Yen Y, Hsieh Y, Kung P, Lan S. Effects of exercise on sleep quality n pregnant women: a systematic review and meta-analysis of randomized controlled trials. *Asian Nurs Res (Korean Soc Nurs Sci)* 2020 Feb;14(1):1-10 [FREE Full text] [doi: [10.1016/j.anr.2020.01.003](https://doi.org/10.1016/j.anr.2020.01.003)] [Medline: [32006719](https://pubmed.ncbi.nlm.nih.gov/32006719/)]
20. Aalami M, Jafarnejad F, ModarresGharavi M. The effects of progressive muscular relaxation and breathing control technique on blood pressure during pregnancy. *Iran J Nurs Midwifery Res* 2016;21(3):331-336 [FREE Full text] [doi: [10.4103/1735-9066.180382](https://doi.org/10.4103/1735-9066.180382)] [Medline: [27186213](https://pubmed.ncbi.nlm.nih.gov/27186213/)]
21. Lawn S, Schoo A. Supporting self-management of chronic health conditions: common approaches. *Patient Educ Couns* 2010 Aug;80(2):205-211. [doi: [10.1016/j.pec.2009.10.006](https://doi.org/10.1016/j.pec.2009.10.006)] [Medline: [19931372](https://pubmed.ncbi.nlm.nih.gov/19931372/)]
22. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [doi: [10.1207/S15324796ABM2601_01](https://doi.org/10.1207/S15324796ABM2601_01)] [Medline: [12867348](https://pubmed.ncbi.nlm.nih.gov/12867348/)]
23. Lenzen SA, Daniëls R, van Bokhoven MA, van der Weijden T, Beurskens A. Disentangling self-management goal setting and action planning: a scoping review. *PLoS One* 2017;12(11):e0188822 [FREE Full text] [doi: [10.1371/journal.pone.0188822](https://doi.org/10.1371/journal.pone.0188822)] [Medline: [29176800](https://pubmed.ncbi.nlm.nih.gov/29176800/)]
24. Bayo-Monton J, Martinez-Millana A, Han W, Fernandez-Llatas C, Sun Y, Traver V. Wearable sensors integrated with Internet of Things for advancing eHealth care. *Sensors (Basel)* 2018 Jun 06;18(6):1851 [FREE Full text] [doi: [10.3390/s18061851](https://doi.org/10.3390/s18061851)] [Medline: [29882790](https://pubmed.ncbi.nlm.nih.gov/29882790/)]
25. de Zambotti M, Rosas L, Colrain IM, Baker FC. The Sleep of the Ring: comparison of the ŌURA sleep tracker against polysomnography. *Behav Sleep Med* 2019;17(2):124-136 [FREE Full text] [doi: [10.1080/15402002.2017.1300587](https://doi.org/10.1080/15402002.2017.1300587)] [Medline: [28323455](https://pubmed.ncbi.nlm.nih.gov/28323455/)]
26. Galea JT, Ramos K, Coit J, Friedman LE, Contreras C, Dueñas M, et al. he use of wearable technology to objectively measure sleep quality and physical activity among pregnant women in Urban Lima, Peru: a pilot feasibility study. *Matern Child Health J* 2020 Jul;24(7):823-828. [doi: [10.1007/s10995-020-02931-5](https://doi.org/10.1007/s10995-020-02931-5)] [Medline: [32350729](https://pubmed.ncbi.nlm.nih.gov/32350729/)]
27. Azimi I, Oti O, Labbaf S, Niela-Vilen H, Axelin A, Dutt N, et al. Personalized maternal sleep quality assessment: An objective IOT-based longitudinal study. *IEEE Access* 2019;7:93433-93447. [doi: [10.1109/access.2019.2927781](https://doi.org/10.1109/access.2019.2927781)]
28. Asgari Mehrabadi M, Azimi I, Sarhaddi F, Axelin A, Niela-Vilén H, Myllyntausta S, et al. Sleep tracking of a commercially available smart ring and smartwatch against medical-grade actigraphy in everyday settings: instrument validation study. *JMIR Mhealth Uhealth* 2020 Nov 02;8(10):e20465 [FREE Full text] [doi: [10.2196/20465](https://doi.org/10.2196/20465)] [Medline: [33038869](https://pubmed.ncbi.nlm.nih.gov/33038869/)]

29. Kinnunen H, Rantanen A, Kenttä T, Koskimäki H. Feasible assessment of recovery and cardiovascular health: accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG. *Physiol Meas* 2020 May 07;41(4):04NT01. [doi: [10.1088/1361-6579/ab840a](https://doi.org/10.1088/1361-6579/ab840a)] [Medline: [32217820](https://pubmed.ncbi.nlm.nih.gov/32217820/)]
30. Henriksen A, Svartdal F, Grimsgaard S, Hartvigsen G, Hopstock L. Polar Vantage and Oura physical activity and sleep trackers: a validation study. *medRxiv* 2020:2020. [doi: [10.1101/2020.04.07.20055756](https://doi.org/10.1101/2020.04.07.20055756)]
31. ACOG Committee Opinion, Number 804: Physical activity and exercise during pregnancy and the postpartum period. *Obstet Gynecol* 2015 Dec;135(4):e178-e188. [doi: [10.1097/aog.0000000000003772](https://doi.org/10.1097/aog.0000000000003772)]
32. Russo MA, Santarelli DM, O'Rourke D. The physiological effects of slow breathing in the healthy human. *Breathe (Sheff)* 2017 Dec;13(4):298-309 [FREE Full text] [doi: [10.1183/20734735.009817](https://doi.org/10.1183/20734735.009817)] [Medline: [29209423](https://pubmed.ncbi.nlm.nih.gov/29209423/)]
33. 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/)]
34. Oura Ring. URL: <https://ouraring.com/> [accessed 2021-10-09]
35. Solivan AE, Xiong X, Harville EW, Buekens P. Measurement of perceived stress among pregnant women: a comparison of two different instruments. *Matern Child Health J* 2015 Sep;19(9):1910-1915 [FREE Full text] [doi: [10.1007/s10995-015-1710-5](https://doi.org/10.1007/s10995-015-1710-5)] [Medline: [25652063](https://pubmed.ncbi.nlm.nih.gov/25652063/)]
36. Austin M, Tully L, Parker G. Examining the relationship between antenatal anxiety and postnatal depression. *J Affect Disord* 2007 Aug;101(1-3):169-174. [doi: [10.1016/j.jad.2006.11.015](https://doi.org/10.1016/j.jad.2006.11.015)] [Medline: [17196663](https://pubmed.ncbi.nlm.nih.gov/17196663/)]
37. Mosack V, Shore ER. Screening for depression among pregnant and postpartum women. *J Community Health Nurs* 2006;23(1):37-47. [doi: [10.1207/s15327655jchn2301_4](https://doi.org/10.1207/s15327655jchn2301_4)] [Medline: [16445363](https://pubmed.ncbi.nlm.nih.gov/16445363/)]
38. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983 Dec;24(4):385-396. [Medline: [6668417](https://pubmed.ncbi.nlm.nih.gov/6668417/)]
39. Roberti JW, Harrington LN, Storch EA. Further psychometric support for the 10-item version of the perceived stress scale. *J Coll Couns* 2006;9(2):135-147. [doi: [10.1002/j.2161-1882.2006.tb00100.x](https://doi.org/10.1002/j.2161-1882.2006.tb00100.x)]
40. Radloff LS. The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. *J Youth Adolesc* 1991 Apr;20(2):149-166. [doi: [10.1007/BF01537606](https://doi.org/10.1007/BF01537606)] [Medline: [24265004](https://pubmed.ncbi.nlm.nih.gov/24265004/)]
41. Neter E, Collins NL, Lobel M, Dunkel-Schetter C. Psychosocial predictors of postpartum depressed mood in socioeconomically disadvantaged women. *Womens Health* 1995;1(1):51-75. [Medline: [9373373](https://pubmed.ncbi.nlm.nih.gov/9373373/)]
42. Mendez DD, Sanders SA, Karimi HA, Gharani P, Rathbun SL, Gary-Webb TL, et al. Understanding pregnancy and postpartum health using ecological momentary assessment and mobile technology: protocol for the Postpartum Mothers Mobile Study. *JMIR Res Protoc* 2019 Jun 26;8(6):e13569 [FREE Full text] [doi: [10.2196/13569](https://doi.org/10.2196/13569)] [Medline: [31244478](https://pubmed.ncbi.nlm.nih.gov/31244478/)]
43. Norman G. Likert scales, levels of measurement and the "laws" of statistics. *Adv Health Sci Educ Theory Pract* 2010 Dec;15(5):625-632. [doi: [10.1007/s10459-010-9222-y](https://doi.org/10.1007/s10459-010-9222-y)] [Medline: [20146096](https://pubmed.ncbi.nlm.nih.gov/20146096/)]
44. 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/)]
45. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008 Apr;62(1):107-115. [doi: [10.1111/j.1365-2648.2007.04569.x](https://doi.org/10.1111/j.1365-2648.2007.04569.x)] [Medline: [18352969](https://pubmed.ncbi.nlm.nih.gov/18352969/)]
46. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 2016 Jul 26;1(3):385-401. [doi: [10.1177/014662167700100306](https://doi.org/10.1177/014662167700100306)]
47. Narita Y, Shinohara H, Kodama H. Resting heart rate variability and the effects of biofeedback intervention in women with low-risk pregnancy and prenatal childbirth fear. *Appl Psychophysiol Biofeedback* 2018 Jun;43(2):113-121. [doi: [10.1007/s10484-018-9389-1](https://doi.org/10.1007/s10484-018-9389-1)] [Medline: [29476282](https://pubmed.ncbi.nlm.nih.gov/29476282/)]
48. Shah Z, Pal P, Pal GK, Papa D, Bharadwaj B. Assessment of the association of heart rate variability and baroreflex sensitivity with depressive symptoms and stress experienced by women in pregnancy. *J Affect Disord* 2020 Dec 01;277:503-509. [doi: [10.1016/j.jad.2020.08.039](https://doi.org/10.1016/j.jad.2020.08.039)] [Medline: [32882507](https://pubmed.ncbi.nlm.nih.gov/32882507/)]
49. Dietz P, Watson ED, Sattler MC, Ruf W, Titze S, van Poppel M. The influence of physical activity during pregnancy on maternal, fetal or infant heart rate variability: a systematic review. *BMC Pregnancy Childbirth* 2016 Oct 26;16(1):326 [FREE Full text] [doi: [10.1186/s12884-016-1121-7](https://doi.org/10.1186/s12884-016-1121-7)] [Medline: [27784276](https://pubmed.ncbi.nlm.nih.gov/27784276/)]
50. Kolovetsiou-Kreiner V, Moertl MG, Papousek I, Schmid-Zaludek K, Lang U, Schlembach D, et al. Maternal cardiovascular and endothelial function from first trimester to postpartum. *PLoS One* 2018;13(5):e0197748 [FREE Full text] [doi: [10.1371/journal.pone.0197748](https://doi.org/10.1371/journal.pone.0197748)] [Medline: [29782509](https://pubmed.ncbi.nlm.nih.gov/29782509/)]
51. Sawane M, Ambade R, Kawthalkar A. Non linear heart rate variability study during pregnancy in Indian women. *JCDR* 2018:1-4. [doi: [10.7860/jcdr/2018/29884.11500](https://doi.org/10.7860/jcdr/2018/29884.11500)]
52. Kotlar B, Gerson E, Petrillo S, Langer A, Tiemeier H. The impact of the COVID-19 pandemic on maternal and perinatal health: a scoping review. *Reprod Health* 2021 Jan 18;18(1):10 [FREE Full text] [doi: [10.1186/s12978-021-01070-6](https://doi.org/10.1186/s12978-021-01070-6)] [Medline: [33461593](https://pubmed.ncbi.nlm.nih.gov/33461593/)]
53. Rhodes A, Kheireddine S, Smith AD. Experiences, attitudes, and needs of users of a pregnancy and parenting app (Baby Buddy) during the COVID-19 pandemic: mixed methods study. *JMIR Mhealth Uhealth* 2020 Dec 09;8(12):e23157 [FREE Full text] [doi: [10.2196/23157](https://doi.org/10.2196/23157)] [Medline: [33264100](https://pubmed.ncbi.nlm.nih.gov/33264100/)]

54. Sedov ID, Anderson NJ, Dhillon AK, Tomfohr-Madsen LM. Insomnia symptoms during pregnancy: a meta-analysis. *J Sleep Res* 2021 Feb;30(1):e13207. [doi: [10.1111/jsr.13207](https://doi.org/10.1111/jsr.13207)] [Medline: [33140514](https://pubmed.ncbi.nlm.nih.gov/33140514/)]
55. Fagherazzi G, El Fatouhi D, Bellicha A, El Gareh A, Affret A, Dow C, et al. An international study on the determinants of poor sleep amongst 15,000 users of connected devices. *J Med Internet Res* 2017 Oct 23;19(10):e363 [FREE Full text] [doi: [10.2196/jmir.7930](https://doi.org/10.2196/jmir.7930)] [Medline: [29061551](https://pubmed.ncbi.nlm.nih.gov/29061551/)]
56. Wilson DL, Barnes M, Ellett L, Permezel M, Jackson M, Crowe SF. Decreased sleep efficiency, increased wake after sleep onset and increased cortical arousals in late pregnancy. *Aust N Z J Obstet Gynaecol* 2011 Feb;51(1):38-46. [doi: [10.1111/j.1479-828X.2010.01252.x](https://doi.org/10.1111/j.1479-828X.2010.01252.x)] [Medline: [21299507](https://pubmed.ncbi.nlm.nih.gov/21299507/)]
57. Tinius R, Polston M, Bradshaw H, Ashley P, Greene A, Parker A. An assessment of mobile applications designed to address physical activity during pregnancy and postpartum. *Int J Exerc Sci* 2021;14(7):382-399 [FREE Full text] [Medline: [34055180](https://pubmed.ncbi.nlm.nih.gov/34055180/)]
58. Kogutt BK, Sheffield JS. Feasibility of remote monitoring to confirm exercise adherence in high risk pregnant women [15N]. *Obstetrics & Gynecology* 2020;135:149S. [doi: [10.1097/01.aog.0000663712.92808.07](https://doi.org/10.1097/01.aog.0000663712.92808.07)]
59. Larsen B, Micucci S, Hartman SJ, Ramos G. easibility and acceptability of a counseling- and mHealth-based physical activity intervention for pregnant women with diabetes: The Fit for Two Pilot Study. *JMIR Mhealth Uhealth* 2020 Oct 21;8(10):e18915 [FREE Full text] [doi: [10.2196/18915](https://doi.org/10.2196/18915)] [Medline: [33084584](https://pubmed.ncbi.nlm.nih.gov/33084584/)]
60. Dodd JM, Louise J, Cramp C, Grivell RM, Moran LJ, Deussen AR. Evaluation of a smartphone nutrition and physical activity application to provide lifestyle advice to pregnant women: The SNAPP randomised trial. *Matern Child Nutr* 2018 Jan;14(1):e12502 [FREE Full text] [doi: [10.1111/mcn.12502](https://doi.org/10.1111/mcn.12502)] [Medline: [28836373](https://pubmed.ncbi.nlm.nih.gov/28836373/)]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists

CES-D: Center for Epidemiologic Studies Depression Scale

EMA: ecological momentary assessment

HRV: heart rate variability

PSS: Perceived Stress Scale

REDCap: Research Electronic Data Capture

REM: rapid eye movement

RMSSD: root mean square of successive difference

STAI: State Trait Anxiety Inventory

S&CC: Smart and Connected Communities

THH: Two Happy Hearts

WIoT: wearable internet of things

Edited by G Eysenbach; submitted 06.06.21; peer-reviewed by B Chaudhry, M Mbwogge; comments to author 11.08.21; revised version received 11.09.21; accepted 18.09.21; published 17.11.21.

Please cite as:

*Jimah T, Borg H, Kehoe P, Pimentel P, Turner A, Labbaf S, Asgari Mehrabadi M, Rahmani AM, Dutt N, Guo Y
A Technology-Based Pregnancy Health and Wellness Intervention (Two Happy Hearts): Case Study*

JMIR Form Res 2021;5(11):e30991

URL: <https://formative.jmir.org/2021/11/e30991>

doi: [10.2196/30991](https://doi.org/10.2196/30991)

PMID: [34787576](https://pubmed.ncbi.nlm.nih.gov/34787576/)

©Tamara Jimah, Holly Borg, Priscilla Kehoe, Pamela Pimentel, Arlene Turner, Sina Labbaf, Milad Asgari Mehrabadi, Amir M. Rahmani, Nikil Dutt, Yuqing Guo. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.11.2021. 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 Smoking Cessation App for Nondaily Smokers (Version 2 of the Smiling Instead of Smoking App): Acceptability and Feasibility Study

Bettina B Hoepfner¹, MS, PhD; Kaitlyn R Siegel¹, BA; Hannah A Carlon², BS; Christopher W Kahler³, PhD; Elyse R Park⁴, MPH, PhD; Susanne S Hoepfner⁵, PhD

¹Recovery Research Institute, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

²Department of Psychology, University of New Mexico, Albuquerque, NM, United States

³Center for Alcohol and Addiction Studies, Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, United States

⁴Health Promotion and Resiliency Research Intervention Program, Psychiatry & Medicine, Massachusetts General Hospital, Boston, MA, United States

⁵Center for Anxiety and Traumatic Stress Disorders, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

Corresponding Author:

Bettina B Hoepfner, MS, PhD

Recovery Research Institute

Department of Psychiatry

Massachusetts General Hospital

151 Merrimac Street

6th Floor

Boston, MA, 02114

United States

Phone: 1 617 643 1988

Fax: 1 617 643 1998

Email: bhoepfner@mgh.harvard.edu

Abstract

Background: Recent evidence highlights the significant detrimental impact of nondaily smoking on health and its disproportionate prevalence in underserved populations; however, little work has been done to develop treatments specifically geared toward quitting nondaily smoking.

Objective: This study aims to test the feasibility, acceptability, and conceptual underpinnings of version 2 of the *Smiling Instead of Smoking* (SiS2) smartphone app, which was developed specifically for nondaily smokers and uses a positive psychology approach.

Methods: In a prospective, single-group study, nondaily smokers (N=100) were prescribed use of the SiS2 app for 7 weeks while undergoing a quit attempt. The app assigned daily positive psychology exercises and behavioral tasks every 2 to 3 days, which guided smokers through using the smoking cessation tools offered in the app. Participants answered surveys at baseline and at 2, 6, 12, and 24 weeks postquit. Feasibility was evaluated based on app use and acceptability based on survey responses. The underlying conceptual framework was tested by examining whether theorized within-person changes occurred from baseline to end of treatment on scales measuring self-efficacy, desire to smoke, and processing of self-relevant health information (ie, pros and cons of smoking, importance of the pros and cons of quitting, and motivation).

Results: Participants used the SiS2 app on an average of 24.7 (SD 13.8) days out of the 49 prescribed days. At the end of treatment, most participants rated the functions of the app as *very easy to use* (eg, 70/95, 74% regarding cigarette log and 59/95, 62% regarding happiness exercises). The average score on the System Usability Scale was 79.8 (SD 17.3; A grade; A+ ≥ 84.1 , B+ < 78.8). Most participants reported that the app helped them in their quit attempt (83/95, 87%), and helped them stay positive while quitting (78/95, 82%). Large effects were found for within-person decreases in the desire to smoke ($b = -1.5$, 95% CI -1.9 to -1.1 ; $P < .001$; $g_{av} = 1.01$), the importance of the pros of smoking ($b = -20.7$, 95% CI -27.2 to -14.3 ; $P < .001$; $g_{av} = 0.83$), and perceived psychoactive benefits of smoking ($b = -0.8$, 95% CI -1.0 to -0.5 ; $P < .001$; $g_{av} = 0.80$). Medium effects were found for increases in self-efficacy for remaining abstinent when encountering internal ($b = 13.1$, 95% CI 7.6 to 18.7; $P < .001$; $g_{av} = 0.53$) and

external ($b=11.2$, 95% CI 6.1 to 16.1; $P<.001$; $g_{av}=0.49$) smoking cues. Smaller effects, contrary to expectations, were found for decreases in motivation to quit smoking ($P=.005$) and the perceived importance of the pros of quitting ($P=.009$). Self-reported 30-day point prevalence abstinence rates were 40%, 56%, and 56% at 6, 12, and 24 weeks after the quit day, respectively.

Conclusions: The SiS2 app was feasible and acceptable, showed promising changes in constructs relevant to smoking cessation, and had high self-reported quit rates by nondaily smokers. The SiS2 app warrants testing in a randomized controlled trial.

(*JMIR Form Res* 2021;5(11):e29760) doi:[10.2196/29760](https://doi.org/10.2196/29760)

KEYWORDS

nondaily; smoking cessation; smartphone app; positive psychology; mHealth; happiness; mobile phone

Introduction

Background

Nondaily smoking continues to be a public health issue with limited empirically supported options to support smoking cessation. Currently, 25.4% of all adult smokers are nondaily smokers [1]. This prevalence has been increasing over the past decade, from a prevalence of 20.2% in 2008 [2] to 25.4% in 2018 [1]. In 2007, the National Institutes of Health highlighted nondaily smoking as a public health issue [3]. Now, more than a decade later, there continues to be a lack of smoking cessation support for this important population of smokers.

For most nondaily smokers, nondaily smoking is a long-term pattern of smoking, not just a transitional phase [4]. Prior research has established that nondaily smokers are more motivated to quit smoking than daily smokers [5]. Moreover, a growing body of research has documented that nondaily smokers differ from daily smokers in terms of numerous characteristics relevant to the process of quitting smoking, including smoking motives [6] and situational antecedents of smoking [7]. These facts argue for the development of targeted and tailored public health efforts to support nondaily smokers in quitting smoking.

Newer emerging evidence has highlighted the urgency of addressing nondaily smoking. Contrary to initial beliefs, the adverse health impact of nondaily smoking is not negligible. Recent research has shown that the mortality risk of native nondaily smokers (ie, those who have never smoked on a daily basis) compared with never smokers is 72% higher [8]. Impacts on all-cause mortality risk have been observed for smoking as few as 6-10 cigarettes per month [9]. These findings clarify that nondaily smoking has a substantial detrimental impact on health. In addition, vulnerable and underserved populations, among whom nondaily smoking is particularly prevalent, are disproportionately affected. These populations include racial and ethnic minority groups [10-12] and persons with mental health and substance use challenges [13]. Thus, addressing nondaily smoking is also a matter of health equity.

To date, very little research has been conducted to develop smoking cessation support approaches geared specifically toward nondaily smokers. Large randomized controlled trials of nicotine replacement therapy (NRT) for nondaily smokers failed to show efficacy in achieving smoking abstinence in comparison with placebo or counseling-only conditions [14,15]. Of note, in these trials, participant adherence to NRT was lower than recommended by study staff, suggesting a lack of interest

among nondaily smokers in NRT as a smoking cessation aid. This observation is in line with other studies that showed that, among college students, nondaily smokers were less likely to be interested in using pharmacotherapy than daily smokers [16] and that nondaily smokers did not view nicotine addiction as relevant to their efforts to quit [17]. This points to the importance of behavioral approaches for nondaily smokers, where mobile technology can play a critical role.

Recent years have seen an uptake in the use [18,19] and clinical linkage of mobile technologies to support smoking cessation, where it is now recommended that health care professionals connect their smoking patients to mobile health (mHealth) resources for smoking cessation [20]. Smartphone app technology may be particularly useful for nondaily smokers, who are less likely to engage in traditional smoking cessation interventions [16]. Beyond treatment modality, the content of intervention approaches for nondaily smokers needs to be tailored to address the unique characteristics of nondaily smoking, as constructs most relevant to smoking cessation for daily smokers appear to be less salient to nondaily smokers as they prepare for the quit day [17].

Development of the Smiling Instead of Smoking App

To address these needs, we developed a smoking cessation app called *Smiling Instead of Smoking* (SiS) [21]. This app builds on the development of positive psychotherapy for smoking cessation [22,23] and adapted its core principles to the app environment. Our rationale for choosing this approach was based on prior research, which indicated that nondaily smokers did not view nicotine addiction as relevant to their efforts to quit but instead viewed themes related to a positive self-identity and wellness as more important [17]. The app's overall conceptual model, as described elsewhere [21,24], is based on social cognitive theory, [25] where goal setting, self-efficacy, and knowledge are leveraged to support behavioral change. Within this framework, the SiS app uses a positive psychology approach to achieve 2 goals: (1) to foster engagement with the app and its smoking cessation content and (2) to support positive affect while smokers undertake the quit attempt. Regarding engagement, positive psychology interventions for health behavior change have been found to be highly appealing to patients [26], resulting in better treatment adherence [27,28] and engagement [29]. Regarding the utility of high positive affect during smoking cessation, findings from laboratory and real time studies have shown that high positive affect is associated with increased self-efficacy [30], decreased desire to smoke [31,32], and greater readiness to process self-relevant

health information [33], all of which are constructs highlighted in dominant theories of health behavior change as playing a causal role in the process of behavioral change, including in the health belief model [34], social cognitive theory [35], the theory of planned behavior [36], the transtheoretical model [37], and the relapse prevention model [38].

We tested the feasibility and acceptability of version 1 of this app (SiS1) in an initial sample of 30 nondaily smokers [24]. Initial testing of this app showed excellent app use, favorable ease of use and usefulness ratings, and significant within-person

changes in line with our conceptual model. Self-reported abstinence rates were also fairly high (53% self-reported abstinence 6 months after quitting) [24].

The SiS2 app was developed based on the results of and feedback from participants using the SiS1 app [24]. In going from the SiS1 to the SiS2 app, we made some overall stylistic changes, such as redesigning the user interface (Figure 1), expanding the variety of daily positive psychology exercises, and adding gamification components (ie, points for completing tasks and a distraction game called *Magma Bear*).

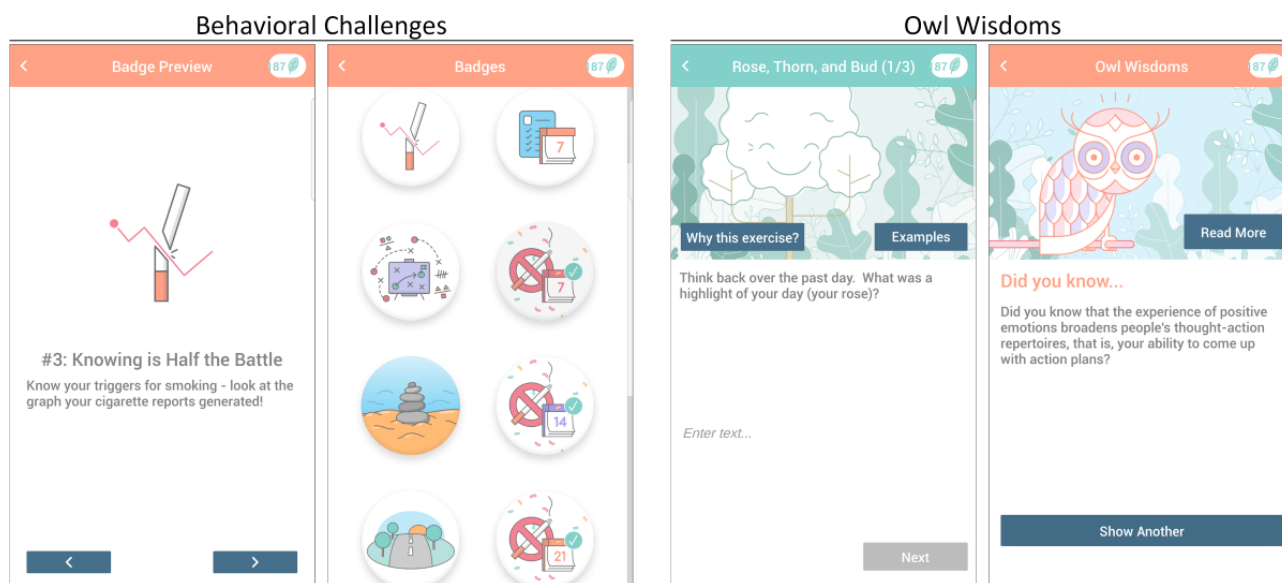
Figure 1. Main menu of Smiling Instead of Smoking app versions 1 and 2. SiS: Smiling Instead of Smoking.



We also made 2 targeted changes to address specific weaknesses. We wanted to make the smoking cessation content more accessible, and we wanted SiS app users to better understand why our positive psychology approach may be useful. To accomplish the former, we replaced our weekly, wordy smoking cessation tutorials that were part of SiS1 with frequent, brief *behavioral challenges* in SiS2. With regard to affecting perceptions about the usefulness of positive psychology

for smoking cessation, we added quite a bit of information on this topic throughout the app. This informational content was optional, and we conceptualized it as an invitation to learn more. To draw attention to this information, SiS2 used a new *Owl Wisdom* mechanism, where the app sent push notifications to app users every 3-4 days to share with them a relevant scientific finding related to happiness or positive affect and its connections to smoking outcomes, health, and well-being (Figure 2).

Figure 2. In going from Smiling Instead of Smoking app version 1 to version 2, behavioral challenges (A) were added to engage app users with the smoking cessation tools and (B) information was added to explain the positive psychology approach via "read more" buttons and Owl Wisdoms.



In study 2, we also changed important parameters regarding the *ask* from participants; that is, instead of onboarding participants in person, as we had done in study 1, we onboarded participants remotely via phone calls in study 2. We also increased the length of the *prescribed* app use from 3 weeks (study 1) to 7 weeks (study 2), as study 1 participants had asked for longer app support. We evaluated the acceptability and feasibility of this new approach in the same manner as in study 1.

Objective

In this paper, we report the outcomes of a study (NCT03951766) examining version 2 of the SiS app (SiS2). Our goals are (1) to assess the feasibility and acceptability of the app as measured by actual app use and ratings of usefulness and ease of use, (2) to test whether within-person changes in line with the underlying conceptual model are observable, and (3) to report smoking cessation outcomes. Regarding the within-person changes, we hypothesize that participants in this study would show, by the end of treatment, an increased self-efficacy to abstain from smoking, reduced desire to smoke, and more positive processing of self-relevant smoking information (eg, decreased perceived pros of smoking and increased perceived cons of smoking). In line with our conceptual model, we do not hypothesize that positive affect would increase during this trial. Rather, the app's goal is to maintain a positive affect while smokers undergo a quit attempt.

Methods

Participants

Participants were adult nondaily smokers who were interested in using a smartphone app to help them quit smoking (recruited from June 11, 2019 to November 15, 2019). Study recruitment information was displayed on Craigslist, Facebook, Reddit, Smokefree.gov, ClinicalTrials.gov, a study recruitment website at the Massachusetts General Hospital, our study website, and websites of local universities. Recruitment was bolstered through word-of-mouth referrals by participants in our first pilot study

and the study staff. To participate, participants had to be aged >18 years, be current nondaily smokers who smoked at least weekly but <25 out of the past 30 days, own an Android or iPhone smartphone, be willing to make a quit attempt as part of the study, be willing to name friends or family members who could help study staff by updating contact information for follow-up assessments, and be fluent in English. The study (NCT03951766) was conducted entirely remotely and was approved by the Mass General Brigham institutional review board. All participants provided informed consent.

Procedure

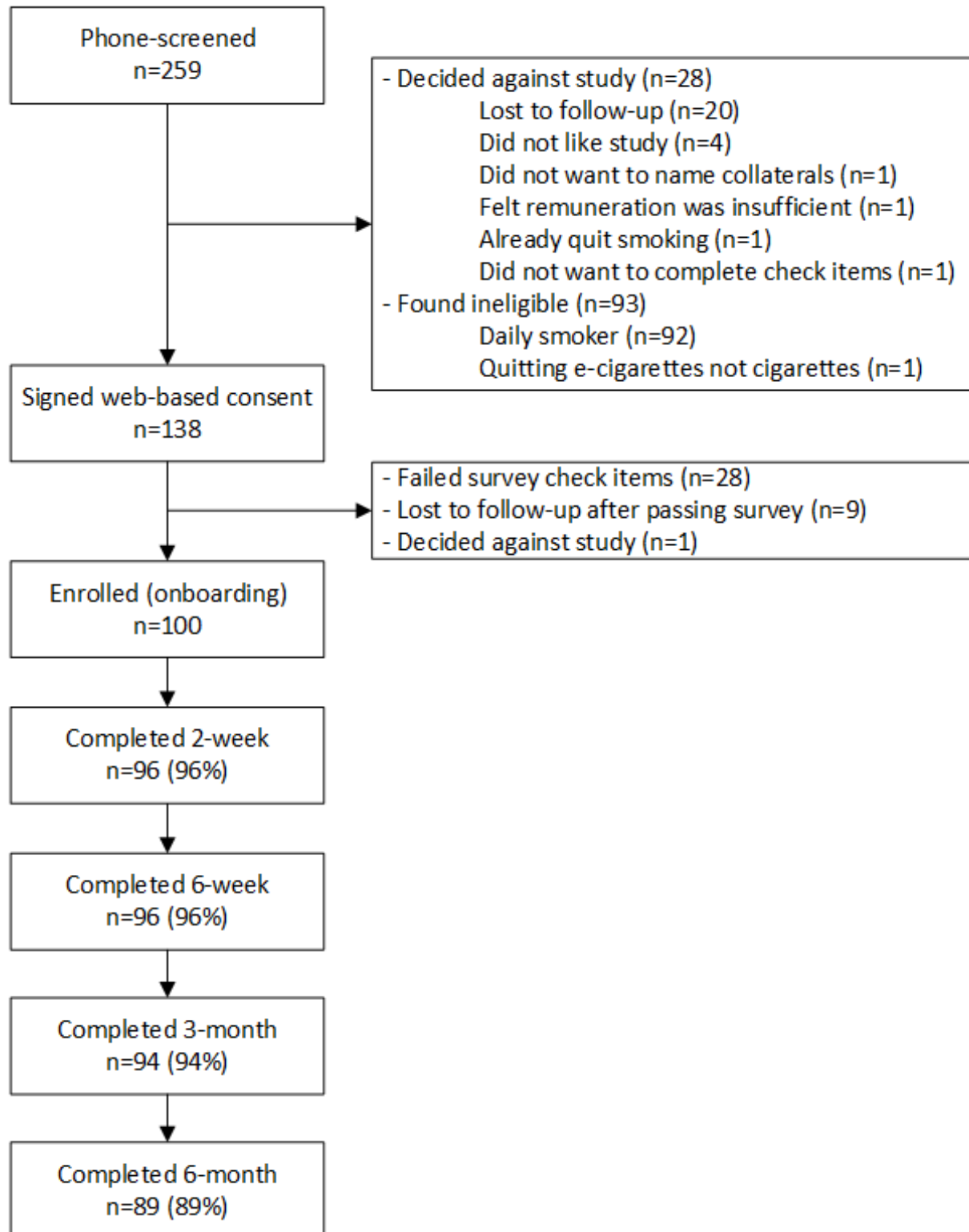
Interested participants were phone-screened and asked to complete a screening test. To pass, participants had to complete a web-based survey and correctly respond to 5 haphazardly placed check questions that tested if respondents were reading survey items. Participants who passed the screening survey were then asked to provide contact information for 2 collaterals who would be able to assist research staff in locating participants for follow-ups, if necessary. Participants were notified by phone if they were eligible, and during this phone call, study staff asked what participants had chosen as their designated quit day to set up the app onboarding call. During the onboarding call, which also served as the study's enrollment call, study staff verified participants' nondaily smoking status and then guided participants through downloading, installing, and using the app. This phone conversation focused on the app and how participants should be using it during the prescribed treatment period. Beyond walking participants through the app, the study staff did not offer further smoking cessation advice. The recording of their app data was confirmed by the study staff during this call. Participants were then asked to use the app for a period of 7 weeks (1 week prequit and 6 weeks postquit) and to complete follow-up surveys on the web 2, 6, 12, and 24 weeks after their initially chosen quit day. Participants received US \$25 for completed surveys or US \$10 for incomplete surveys or surveys with failed check items. They received US \$50 for the week 6 survey (end of treatment), which was longer than

the other surveys. Participants provided their social security numbers to enable remuneration by check. All surveys were administered via the electronic data capture system REDCap (Research Electronic Data Capture) [39].

In total, we phone-screened 259 individuals (Figure 3). Of these 259 individuals, 93 (35.9%) were not eligible (of the 93 individuals, 92 [98.9%] were not eligible as they were daily

smokers and 1 [1.1%] were not eligible as they wanted to quit vaping), and 28 (10.8%) decided against the study. The remaining 53.5% (138/259) signed the web-based consent and started the screening test. Of these 138 individuals, 28 (20.3%) failed the check items embedded in the survey, and 10 (7.2%) passed but decided against the study at this point. The remaining 72.5% (100/138) of individuals were onboarded to SiS2.

Figure 3. Consolidated Standards of Reporting Trials diagram for Smiling Instead of Smoking app study 2.



Participants who missed scheduled assessments were contacted with multiple reminders and were contacted for subsequent assessments unless they actively withdrew from the study. Survey completion rates were 96%, 96%, 94%, and 89% for follow-up surveys 2, 6, 12, and 24 weeks after the chosen smoking quit date, respectively (N=100). For half of the participants (49/100, 49%), the 24-week follow-up occurred after the onset of the COVID-19 pandemic in the United States.

Treatment

Participants received the SiS2 app, which was based on existing, empirically tested positive psychology exercises [40-44], published guidelines for smoking cessation (ie, US Clinical Practice Guidelines) [45], and user feedback from nondaily smokers who used version 1 of the app [24]. The app engaged participants in daily positive psychology exercises over the course of 7 weeks, as well as temporally appropriate *behavioral challenges* every 2-3 days that were designed to engage participants with ad libitum tools offered by the app. The SiS2 app used 2 new happiness exercises in addition to the 3 exercises used in the SiS1 app (ie, 3 *Good Things*, where participants entered text describing 3 good things that happened to them that day; *Savoring*, where participants entered text describing 2 experiences they savored, and *Experiencing Kindness*, where participants described an act of kindness they did or one they witnessed). The new exercises were *Rose, Thorn, and Bud*, where participants described one good and one challenging thing in the past 24 hours and one thing they looked forward to in the next 24 hours, and *Reliving Happy Moments*, where participants viewed and then described a picture they had taken that made them happy [41]. One of these five exercises was chosen at random each day by the app to be completed that day. The ad libitum tools offered tracking functionality (ie, logging smoked cigarettes), graphical summaries (ie, pie chart of reasons smoked with matching strategies to stay smoke free in such situations), reminders (ie, a tool to send push notifications to stay smoke free at specific times), note keeping (ie, personal reasons for quitting), and information (eg, benefits of quitting and strategies). The SiS2 app used push notifications in a variety of ways to engage app users (eg, reminders for missed exercises and push notification for new *behavioral challenges* and *Owl Wisdoms*). The design of the SiS2 app is evidence based and in line with all 6 published recommendations for smoking cessation apps [46]: it (1) is available at no cost, (2) keeps information private (ie, saved on our hospital servers), (3) matches individual needs and interests (eg, suggested strategies are matched to logged smoking reasons), (4) adapts as one's needs and interests change (eg, behavioral challenges are anchored on the resettable quit day to be temporally appropriate), (5) helps to manage nicotine withdrawal symptoms (eg, specific strategies are offered in the app), and (6) allows users to track progress (eg, keeps track of days since quit day, provides graphs over time, and awards badges for achieving milestones).

Measures

Baseline Characteristics

During the web-based baseline survey, sociodemographic information was collected, including age, gender, race, and education level. Participants also answered questions about their

smoking characteristics (ie, number of days smoked in the past 30 days, number of cigarettes smoked per smoking day, history of daily smoking—yes or no—and past quit attempts).

Measures of Feasibility and Acceptability

App Use

App use was automatically recorded by the app, which time stamped every interaction with the app. From these data, we calculated the number of days on which participants used the app during the *prescribed* period of app use (ie, 0-49 days) and examined the number of days per week participants used the SiS2 app.

System Usability Scale

App usability was measured at the end of treatment using the 10-item System Usability Scale (SUS) [47], where we used the phrase *SiS app* in place of *this system*. Participants rated items (eg, "I thought the SiS app was easy to use.") on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). Scores are interpreted as *grades*: A+ for scores ≥ 84.1 , A for scores 84.1-78.8, B+ for scores <78.8; scores <70 indicate *below average* usability [48].

App's Ease of Use and Usefulness

At the end of treatment (week 6), participants rated the ease of use and usefulness of each component of the app (17 items each, listed in the *Results* section) on a 4-point Likert scale, where ease of use was rated as 0=not easy at all, 1=somewhat easy to use, 2=easy to use, or 3=very easy to use, and usefulness was rated as 0=not at all useful, 1=somewhat useful, 2=useful, or 3=very useful.

Perceptions of How the App Might Have Helped

At the end of treatment (week 6), participants provided their level of agreement (5-point Likert scale, 1=strongly disagree to 5=strongly agree) on 17 items pertaining to the helpfulness of the app during the quitting process (eg, in preparing for the quit attempt and during risky situations). Participants also indicated whether the app helped them in their quit attempt and if they would recommend it to a friend who wanted to quit smoking (yes or no).

Measures of Constructs Hypothesized to Change Over Time

As part of all surveys, participants completed the following scales:

Self-Efficacy

The Smoking Self-Efficacy Questionnaire (12-items; slider scale; 0=not at all confident to 100=extremely confident) [49] assesses confidence to abstain from smoking when faced with internal stimuli (eg, "when I feel very anxious") and external stimuli (eg, "after a meal").

Desire to Smoke

The Brief Questionnaire of Smoking Urges [50] uses 10 items on a 7-point Likert scale (1=strongly disagree to 7=strongly agree) to measure one's desire to smoke as it relates to reward (eg, "A cigarette would taste good now") and relief from

negative affect (eg, “I could control things better right now if I could smoke”).

Processing Self-Relevating Health Information

The Attitudes Towards Smoking Scale [51] assesses participants’ feelings toward adverse effects, psychoactive benefits, and pleasures of smoking. Participants rated 18-items (eg, “smoking is ruining my health”) on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). The Decisional Balance Inventory for Smoking–Short Form (6-items; slider scale; 0=not at all important to 100=extremely important) [52] assesses how expectations that are positive (eg, “Smoking cigarettes relieves tension.”) and negative (eg, “I’m embarrassed to have to smoke”) weigh in on one’s decision to smoke at that moment (ie, “right now”). The impact of perceived benefits and barriers on quitting smoking was evaluated using 2 single-item measures: (1) “Think about all the things you LIKE/LOVE about quitting/being smoke-free; taken together, how important are those things to you RIGHT NOW?” and (2) “Think about all the things you DISLIKE/HATE about quitting/being smoke-free; taken together, how important are those things to you RIGHT NOW?”; both were rated on slider scales ranging from 0=not at all to 100=extremely important. The Commitment to Quitting Smoking Scale [53] asks participants to rate their level of agreement (Likert scale, 1=strongly disagree to 5=strongly agree) on 8-items assessing motivation to quit smoking (eg, “I’m not going to let anything get in the way of my quitting smoking”). In addition, a single-item slider scale (0=not at all to 100=extremely motivated) directly asked participants, “How MOTIVATED are you to quit smoking/stay quit?”

Positive Affect

The Positive and Negative Affect Schedule [54] uses 20 items (10 positive adjectives and 10 negative adjectives) on a 5-point Likert scale (1=very slightly or not at all to 5=extremely) to assess the extent to which participants experienced positive and negative affect in the past week. Participants also answered 2 questions using a single-item slider (0=not at all to 100=extremely) to indicate their happiness in the moment and over the past week. Overall life-satisfaction and happiness were assessed using the Satisfaction with Life Scale (5 items; Likert scale ratings from 1=strongly disagree to 7=strongly agree) [55] and the Subjective Happiness Scale (4 items; item-specific anchor points; eg, for “Compared with most of my peers, I consider myself [...],” ratings range from 1=less happy to 7=more happy) [56].

Exploratory Outcome—Self-reported Abstinence

During each survey, participants were asked to rate their smoking status using the following options: “I smoke daily,” “I smoke nondaily (and have smoked in the past 7 days),” “I smoke nondaily (but have NOT smoked in the past 7 days),” and

“I do not smoke at all.” Participants who reported not smoking at all were then asked if they had been completely abstinent since their originally chosen quit day (if no, then since when), during the past 7 days, and during the past 30 days.

We did not perform biochemical tests to confirm self-reported abstinence in line with guidance (at that time) that such tests should not be used for studies with no face-to-face contact and studies in which data are optimally collected through the internet, telephone, or mail [57].

Analytic Strategy

We used SAS 9.4 for all analyses. To describe feasibility, acceptability, and smoking cessation outcomes, we calculated descriptive statistics. To test if theorized within-person changes occurred from baseline to end of treatment, we fit one repeated measures mixed effects model per construct hypothesized to change over time (ie, self-efficacy, desire to smoke, and processing of self-relevant health information). In these models, the sole predictor was time, called TIME, modeled categorically (ie, baseline, 2, 6, 12, and 24 weeks). Observations were modeled as nested within individuals using an unstructured covariance matrix. Per protocol, the primary end point of interest was end of treatment (ie, 6 weeks after the chosen quit date). Thus, using this model, we reported the pairwise contrast between baseline and week 6. Given the exploratory nature of this study, we did not correct for multiple testing. Effect sizes for within-person changes from baseline to week 6 in each outcome measure were reported as Hedges g_{av} , a bias-corrected effect size estimate recommended for correlated samples [58]. We used the same modeling approach to capture changes in positive affect. For the 2 outcomes that indicated significant effects counter to the hypothesized direction (ie, motivation to quit and the perceived importance of the pros of quitting), we used post hoc analyses to see whether changes from baseline to end of treatment differed between participants who achieved 30-day point prevalence abstinence at the end of treatment versus those who did not. In these post hoc models, we added quit status as a binary predictor (QUIT: 1=abstinent, 0=not) and its interaction with TIME to the model. We interpreted a significant QUIT*TIME effect as an evidence of differential changes over time based on quit status.

Results

Participant Characteristics

Participants were predominantly (70/100, 70%) converted nondaily smokers (ie, people who had smoked daily previously), who smoked an average of 4.6 (SD 3.3) cigarettes per smoking day on 14.7 (SD 4.6) days out of the past 30 days, in line with expected rates of smoking in nondaily smokers [59]. Many had made a previous quit attempt, also in line with expectations for this type of smoker [5]. Demographics are reported in Table 1.

Table 1. Demographics and smoking characteristics (N=100).

Characteristics	Values
Demographics	
Age (years), mean (SD)	35.9 (11.4)
Gender (female), n (%)	61 (61)
Race, n (%)	
White	75 (75)
Black	14 (14)
Other or unknown	11 (11)
Hispanic	12 (12)
Education, n (%)	
High school or less	15 (15)
Some college	47 (47)
BA ^a , BS ^b , or higher	38 (38)
Smoking characteristics	
Number of days smoked in past 30 days, mean (SD)	14.7 (4.6)
Number of cigs smoked per smoking day, mean (SD)	4.6 (3.3)
Ever smoked daily? (yes), n (%)	70 (70)
Ever quit before? (yes), n (%)	77 (77)

^aBA: Bachelor of Arts.

^bBS: Bachelor of Science.

Feasibility as Captured by App Use

Participants used the SiS2 app on average on 24.1 (SD 14.1, range 0-49) days out of the 49 prescribed days after onboarding (49% of days). Most participants used the app on their own on the day after their onboarding call (96/100, 96%). During the initial week of app use (ie, the week leading up to the chosen quit day), the percentage of participants interacting with the app on a given day decreased to 60% (60/100). This decrease was largely because of participants settling into a less-than-daily routine of using the app rather than participants discontinuing app use altogether: 7% (7/100) used the app on 0-1 days after onboarding during the first week, 53% (53/100) on 2-5 days, and 40% (40/100) on 6-7 days. App use declined further at a gradual pace during the remaining 6 weeks of prescribed app use, so that during the last week, 33% (33/100) used the app on 0 days, 12% (12/100) on 1 day, 35% (35/100) on 2-5 days, and

20% (20/100) on 6-7 days. More than a quarter of the participants (28/100, 28%) used the SiS2 app at least three times per week for every week of treatment.

Acceptability as Captured by the End of the Treatment (Week 6) Survey Responses

The average score on the SUS, as measured at the end of treatment, was 79.8 (SD 17.3), which represents an A grade [48]. Participants rated the specific functions of the app (Table 2) as *easy to very easy* to use. These ratings ranged from 2.3 (SD 0.9) for playing *Magma Bear* to 2.7 (SD 0.6) for viewing earned badges. Essential features, including completing the positive psychology exercises, completing the behavioral challenges, and using the cigarette log, all scored high and in this range, with 62% (59/95), 64% (61/95), and 74% (70/95), indicating that these functions, respectively, were *very easy to use*.

Table 2. User ratings of the Smiling Instead of Smoking app version 2 functions (N=95).

User ratings	Ease of use ^a		Useful ^b	
	Value, mean (SD)	Very easy, n (%)	Value, mean (SD)	Very useful, n (%)
System Usability Scale				
SUS ^c score ^d	79.8 (17.3)	— ^e	—	—
Happiness related tasks				
Completing the positive psychology exercises every day	2.4 (0.9)	59 (62)	2.2 (0.9)	44 (47)
Specifically, completing <i>3 Good Things</i>	2.4 (0.9)	55 (58)	2.2 (0.9)	48 (51)
Specifically, completing <i>Savoring</i>	2.4 (0.8)	56 (59)	2.2 (0.9)	45 (47)
Specifically, completing <i>Experiencing Kindness</i>	2.4 (0.9)	58 (62)	2.1 (1.0)	45 (47)
Specifically, completing <i>Reliving Happy Moments</i>	2.5 (0.8)	64 (67)	2.2 (0.9)	47 (49)
Specifically, completing <i>Rose, Thorn, and Bud</i>	2.4 (0.8)	56 (59)	2.1 (1.0)	43 (45)
Viewing the Happiness Log of past exercise completions.	2.6 (0.7)	70 (74)	2.2 (0.9)	42 (44)
Viewing Owl Wisdoms (ie, happiness science findings)	2.6 (0.6)	61 (64)	2.2 (0.9)	45 (47)
Smoking specific tasks				
Setting (and if applicable resetting) the quit day	2.5 (0.8)	62 (65)	2.1 (1.0)	42 (44)
Completing the behavioral challenges	2.5 (0.7)	61 (64)	2.2 (0.9)	46 (48)
Accessing and updating the cigarette log	2.6 (0.7)	70 (74)	2.4 (0.8)	53 (56)
Using the <i>Magma Bear</i> game	2.3 (0.9)	54 (57)	1.3 (1.2)	26 (27)
Setting and receiving <i>Smoke Alarms</i>	2.4 (0.8)	53 (56)	2.1 (1.0)	42 (44)
Specifying personal reasons for quitting smoking	2.5 (0.8)	59 (62)	2.3 (0.8)	51 (54)
Viewing strategies for remaining smoke-free	2.5 (0.7)	60 (63)	2.3 (0.9)	49 (52)
Viewing earned badges	2.7 (0.6)	66 (69)	2.0 (1.0)	40 (43)
Viewing benefits of quitting smoking	2.6 (0.6)	61 (64)	2.2 (0.9)	45 (48)

^aEase of use was rated on a 4-point scale: 0=not easy at all, 1=somewhat easy to use, 2=easy to use, 3=very easy to use.

^bUsefulness was rated on a 4-point scale: 0=not at all useful, 1=somewhat useful, 2=useful, 3=very useful.

^cSUS: System Usability Scale; scores can range from 0 (very poor perceived usability) to 100 (excellent perceived usability) in 2.5-point increments.

^dA+=84.1-100; A=80.3-84.0; B=68-80.3.

^eThe System Usability Scale (SUS) presents a general usability score (does not differentiate between ease of use and usefulness) and uses a different rating scale (1=strongly disagree, 5=strongly agree) than the app's ease of use and usefulness scale (ease of use was rated as 0=not easy at all, 1=somewhat easy to use, 2=easy to use, 3=very easy to use, and usefulness was rated as 0=not at all useful, 1=somewhat useful, 2=useful, 3=very useful). For this reason, SUS values for ease of use (very easy) and useful (mean, very useful) were not included.

Greater variations existed in the perceived usefulness of the functions. These ratings ranged from 1.3 (SD 1.2) for playing *Magma Bear* to 2.4 (SD 0.8) for using the cigarette log. Other than *Magma Bear*, all functions scored as at least *useful* or higher on average. In particular, high-scoring functions were the cigarette log (53/95, 56% of participants rated it as *very useful*), viewing strategies for remaining smoke-free (49/95, 52%), and the happiness exercise *Reliving Happy Moments* (47/95, 49%).

In rating how the SiS2 app might have helped them to quit smoking (Table 3), nearly all participants (87/95, 92% of participants who responded to this item) indicated that the SiS2 app served as a useful reminder of why quitting smoking was

important to them. The next most useful app aspect, by participant ratings, was that the SiS2 app showed them how happiness was important in numerous ways (82/95, 86% of participants). Organizational and teaching goals were also met by the SiS2 app, with participants indicating that the SiS2 app helped them stay on track (80/95, 84% of participants), helped them prepare for the quit attempt (80/95, 84% of participants), and provided confidence in the steps to take (78/95, 82% of participants). For some smokers, the SiS2 app was useful in dealing with specific risky smoking times that arose during this quit attempt (65/95, 68% of participants). Overall, most participants felt that the SiS2 app helped them in their quit attempt (83/95, 87% of participants) and that most would recommend it to a friend (87/95, 92% of participants).

Table 3. Ratings of how the Smiling Instead of Smoking app version 2 might have helped (N=95).

User ratings ^a	Value, mean (SD)	Agree and Strongly agree, n (%)
The SiS^b app...		
...reminded me why I wanted to quit.	4.4 (0.9)	87 (92)
...reminded me that quitting was important to me.	4.4 (0.8)	87 (92)
...showed me how happiness is important in numerous ways.	4.3 (0.9)	82 (86)
...made me think that it was worthwhile for me to quit.	4.3 (0.9)	81 (85)
...helped remind me to stay on track with quitting.	4.2 (1.0)	80 (84)
...helped me prepare for the quit attempt.	4.3 (1.0)	80 (84)
...helped me stay positive while quitting.	4.2 (1.0)	78 (82)
...gave me a sense of accomplishment as I progressed through my quit attempt.	4.3 (0.9)	78 (82)
...made me feel that I knew the right steps to take to quit.	4.2 (0.9)	78 (82)
...made me take my quit attempt seriously.	4.2 (1.0)	76 (80)
...gave me confidence that I can quit smoking.	4.0 (0.9)	73 (77)
...reminded me in crucial moments to stay quit.	4.1 (1.0)	72 (76)
...made me feel that someone cared if I quit.	4.1 (1.0)	70 (74)
...encouraged me when things were getting tough.	4.0 (1.0)	70 (74)
...gave me the feeling I could get trusted advice at any time.	3.9 (1.1)	67 (71)
...helped me deal with risky smoking times.	3.9 (1.1)	65 (68)
...gave me a new toy to play with rather than dwell on quitting.	3.7 (1.3)	60 (63)
Overall ratings		
Would you recommend the SiS app to a friend who wants to quit smoking? (yes)	— ^c	87 (92)
Taken altogether, do you think that the app helped you in your quit attempt? (yes)	—	83 (87)

^aRated on a 5-point Likert scale: 1=strongly disagree, 2=disagree, 3=neither agree nor disagree, 4=agree, and 5=strongly agree.

^bSiS: Smiling Instead of Smoking.

^cThe overall rating questions were answered as yes or no, therefore, there is no mean for this item.

Tests of Within-Person Changes Predicted by the SiS2 Conceptual Model

Most, but not all, of the hypothesized within-person changes were observed from baseline to end of treatment (Table 4). In line with our conceptual model, the results indicated that the self-efficacy for remaining abstinent significantly increased for both internal cues ($P<.001$) and external cues ($P<.001$). Furthermore, participants' desire to smoke decreased ($P<.001$)

and perceptions of smoking became less positive, as expressed through reduced valuing of psychoactive benefits and pleasure of smoking, as well as reductions in the perceived importance of the pros of smoking (Table 4). Effects were large for decreases in desire to smoke, importance of the pros of smoking, and ratings of the psychoactive benefits of smoking (ie, all Hedges $g_{av} \geq 0.80$), and effects were of medium size for self-efficacy and the pleasure of smoking (ie, all Hedges $g_{av} \geq 0.49$).

Table 4. Within-person changes on theorized mechanisms of change from baseline to end of treatment.

Construct (scale)	Cronbach α at baseline ^a	Baseline ^b , mean (SD)	Scale range	6-week ^c versus baseline		
				b ^d (95% CI)	P value	g _{av}
Self-efficacy						
SEQ-12 ^e (internal cues)	.86	53.4 (22.0)	0-100	13.1 (7.6 to 18.7)	<.001	0.53
SEQ-12 (external cues)	.77	58.9 (21.1)	0-100	11.1 (6.1 to 16.1)	<.001	0.49
Desire to smoke						
QSU ^f (smoking urges)	.92	3.7 (1.4)	1-7	-1.5 (-1.9 to -1.1)	<.001	1.01
Processing self-relevant health information						
Positive appraisals of smoking						
ATS ^g (psychoactive benefits)	.77	4.0 (0.7)	1-5	-0.8 (-1.0 to -0.5)	<.001	0.80
ATS (pleasure)	.84	3.3 (1.0)	1-5	-0.6 (-0.8 to -0.3)	<.001	0.52
DCB ^h (importance of the pros of smoking)	.67	56.6 (19.9)	0-100	-20.7 (-27.2 to -14.3)	<.001	0.83
Negative appraisals of smoking						
ATS (adverse effects)	.84	4.4 (0.5)	1-5	0.0 (-0.1 to 0.2)	.76	0.03
DCB (importance of cons of smoking)	.80	68.6 (26.2)	0-100	-2.9 (-8.1 to 2.3)	.27	0.09
Benefits and barriers to quitting smoking						
Single item (pros of quitting)	N/A ⁱ	84.9 (21.1)	0-100	-9.1 (-15.9 to -2.3)	.009	0.35
Single item (cons of quitting)	N/A	63.4 (32.3)	0-100	-5.1 (-13.7 to 3.6)	.25	0.14
Motivation to quit smoking						
CQSS ^j (commitment to quitting)	.89	4.1 (0.7)	1-5	-0.1 (-0.3 to 0.1)	.25	0.09
Single item (how motivated)	N/A	88.0 (14.6)	0-100	-6.6 (-11.1 to -2.0)	.005	0.34
Positive affect						
PANAS ^k (past week positive affect)	.66	3.0 (0.6)	1-5	0.0 (-0.1 to 0.2)	.45	0.08
PANAS (past week negative affect)	.72	2.7 (0.7)	1-5	0.0 (-0.1 to 0.2)	.74	0.02
Single item (how happy past week)	N/A	67.7 (21.5)	0-100	1.2 (-3.8 to 6.3)	.62	0.09
Single item (how happy right now)	N/A	70.0 (20.4)	0-100	-0.5 (-5.4 to 4.4)	.83	0.00
Satisfaction with life	.87	4.4 (1.4)	1-7	0.3 (-0.1 to 0.6)	.11	0.18
Subjective happiness	.87	4.8 (1.4)	1-7	0.2 (-0.1 to 0.4)	.16	0.12

^aCronbach α is a measure of the internal consistency of each scale at baseline.

^bBaseline occurred before Smiling Instead of Smoking app download.

^c6-week follow-up occurred at the end of the prescribed 49 days of app use (ie, 6 weeks post quit day).

^db is the parameter estimate of the pairwise difference of week 6 compared with baseline from the repeated measures mixed effects model.

^eSEQ-12: Smoking Self-Efficacy Questionnaire.

^fQSU: Brief Questionnaire of Smoking Urges.

^gATS: Attitudes Towards Smoking Scale.

^hDCB: Decisional Balance Inventory for Smoking.

ⁱN/A: not applicable (Cronbach α is not applicable to single-item measures).

^jCQSS: Commitment to Quitting Smoking Scale.

^kPANAS: Positive and Negative Affect Schedule.

Contrary to expectations, motivation to quit smoking and perceived importance of the pros of quitting decreased below baseline levels by the end of treatment ($P=.005$ and $P=.009$, respectively). Both effects were smaller than the other observed

effects (ie, Hedges $g_{av} \leq 0.35$). Post hoc analyses showed that motivation to quit decreased from baseline to week 6 for those who did not succeed in quitting ($b=-14.7$, 95% CI -20.1 to -9.2 ; $P<.001$), whereas it tended to increase for those who did

($b=5.0$, 95% CI -1.5 to 11.4 ; $P=.13$; interaction $P<.001$). No such interaction effect existed for the perceived importance of pros of quitting (abstinent: $b=-9.5$, 95% CI -20.0 to 1.1 ; $P=.08$; nonabstinent: $b=-8.9$, 95% CI -17.8 to 0.0 ; $P=.05$; interaction $P=.61$). We were unable to detect any changes in the negative appraisals of smoking, cons of quitting, or commitments to quitting. Of note, the perceptions of the adverse effects of smoking (ie, Attitudes Towards Smoking Adverse Effects subscale) were already very high at baseline, with an average score of 4.4 (SD 0.5) out of 5 (Table 4), indicating the possibility of a ceiling effect for this measure. However, in terms of the importance of these cons of smoking, there was some room for improvement, with an average baseline level of 68.8 (SD 26.2) on a scale of $0-100$. The results of the exploratory analyses of changes in positive affect are included in Table 4. In line with our expectations, we were unable to detect any changes in positive affect while smokers underwent the quit attempt.

Self-reported Smoking Cessation Rates

The self-reported 30-day point prevalence abstinence rates at the end of treatment (ie, 6 weeks after the quit date) was 40% ($40/100$ participants; participants who did not complete the survey were presumed to be smoking). Self-reported abstinence increased further during follow-up, with 30-day point prevalence abstinence reported by 56% ($56/100$) of participants at both 12 and 24 weeks after the quit day. Participants who did not complete the surveys were assumed to be smoking.

Discussion

Principal Findings

We tested version 2 of the SiS app (SiS2) in a single-group longitudinal study using minimal contact and no face-to-face contact. We found that the SiS2 app was a feasible and acceptable mHealth tool to support nondaily smokers in quitting smoking.

Nondaily smokers were engaged well by the SiS2 app ($67/100$, 67% of participants still used the app in the seventh week after enrollment) and found it useful. App use is an important outcome metric as it is through interaction with an app that app users are engaged in therapeutic activities. Achieving high app use is currently a critical barrier that challenges the utility of publicly available apps, and a recent study on mental health apps found that $<4\%$ of app users used the app 15 days after installation in real-world settings [60]. In clinical trials, app use is expected to be higher for a number of reasons (eg, actively procuring commitment from study participants to study procedures, which they may interpret to include app adherence) [61] but is nevertheless a pivotal marker of feasibility. In our study, participants used the SiS2 app for an average of 24.7 days. To put this app use into context, it may be useful to contrast it with app use observed in a recent randomized controlled trial [62], which represents the current best evidence regarding smoking cessation apps. In this study, Bricker et al [62] used a double-blind, individually randomized, 2-group stratified design to test the efficacy of their app, called iCanQuit, which is based on acceptance and commitment therapy. They compared it with the publicly available smartphone app QuitGuide, which was developed and disseminated by the

National Cancer Institute (NCI). In this study, participants used the QuitGuide app on 7.1 days, on average, and the iCanQuit app on 24.3 days. Thus, the app use of 24.7 days observed for the SiS2 app is in line with the best standard set to date by the iCanQuit app. However, in our study, participants engaged in a phone call with study staff to get oriented to the app, in line with how we conceptualized such an app should be used in a health care setting, which was not the case for the Bricker et al [62] study, in which participants only received a link to download and install the app [62]. We also told participants at baseline that we would like to conduct a Skype (Microsoft Corporation) interview with a subsample of the study participants ($20/100$, 20%) at the end of the treatment about their experience using the app and asked them if they would like to be considered for such an interview (US \$25 was offered for the Skype interview; $95/100$, 95% said yes). Although it is standard practice for this type of treatment evaluation research to include an exit interview, this question may have also created motivation for greater app use than may happen in a real-world setting, although notably, following up with patients after a smoking cessation referral is part of best practice guidelines (ie, the *arrange follow-up* portion of the 5As) [63]. In line with the *law of attrition* for digital technologies [64], there was app use attrition after an initial curiosity phase; however, there was also sustained use of the app throughout the 7 weeks of the prescribed app use period, and more than a quarter of our sample ($28/100$, 28%) engaged meaningfully with the SiS2 app every week of the 7-week treatment period. This finding was particularly encouraging, given that we had increased the prescribed app use period from 3 to 7 weeks in going from version 1 to version 2 of the app.

In terms of the perceived usefulness of the app, our data further suggest that we were able to improve perceptions of the usefulness of the positive psychology content of the SiS app. In comparing usefulness ratings across the SiS1 and SiS2 studies, we noted that the perceived usefulness of *completing the positive psychology exercises every day* increased from 1.8 in study 1 [24] to 2.2 in study 2. It is difficult, of course, to draw direct conclusions from this improvement in scores as the demographics of the samples were also quite different (ie, largely Black males in SiS1 vs White females in SiS2). Overall, 87% ($83/95$) of the nondaily smokers participating in our study felt that using the SiS2 app helped them in their quit attempt. This compares favorably with the usefulness ratings of 80% for the iCanQuit app and 72% for the NCI QuitGuide app [62]. Similarly, in terms of recommending the app to a friend, SiS2 fared well ($87/95$, 92% for SiS2 compared with 83% and 71% for the iCanQuit and QuitGuide apps, respectively) [62]. Together, these findings demonstrate high levels of engagement and positive perceptions of the usefulness of the SiS2 app.

In terms of the user interface, SiS2 scored well (A grade, 79.8) but left room for improvement. To put this score into context, it may be useful to consider 2 other studies that have used the SUS scale to evaluate smoking cessation apps. In an early stage study of user experiences of smokers with serious mental illness ($n=5$), the NCI smartphone app *QuitPal* received a SUS score of 66 [65]. In a larger trial in the same study population, participants were randomized to use the NCI app *QuitGuide*

or the investigator developed app *Learn to Quit* [66]. In this study, SUS scores were 78 and 85, respectively, for the apps *QuitGuide* and *Learn to Quit*. In these 2 studies, onboarding was performed in person and included additional help over time, as needed. In our case, onboarding was performed remotely in a single phone call. Thus, our robust SUS score suggests that the SiS2 app is implementable via low- to no-touch linkage approaches. Our data further suggest that the change from weekly smoking cessation modules to shorter, more frequent *behavioral challenges* improved the ease of use of the app. In comparing the SiS1 and SiS2 ease of use scores for these features of the apps, we saw a noticeable increase: from a score of 2.0 (*easy to use*) for *completing the smoking sessions* [24] to 2.5 (midpoint between *easy to use* and *very easy to use*) for *completing the behavioral challenges*. This suggests that it may indeed be useful to break larger asks within an app into smaller, simpler asks, so long as they are meaningful.

Findings from the within-person tests were largely in line with our conceptual model. Overall perceptions regarding smoking changed as expected: participants reported greater confidence to abstain from smoking when faced with internal or external stimuli, experienced a lower desire to smoke, and decreased their positive appraisal of smoking over time. Notable exceptions to our expected within-person changes were the unexpected decreases in the single-item measures capturing motivation to quit smoking and the importance of the pros of quitting. Our post hoc analyses provide some insight into possible explanations for the decrease in motivation, where it appeared that motivation to quit smoking was negatively affected by having tried and failed, in line with the abstinence violation effect [67]. The decrease in the perceived importance of the pros of quitting smoking suggests that as smokers enter the maintenance phase of smoking cessation, the pros of quitting have decreased salience. As these scores at this point in time do not relate to smoking cessation success prospectively, this finding suggests that it may not be fruitful to emphasize the pros of quitting at this point of the smoking cessation process in mHealth tools designed to support quitting smoking and maintaining smoking abstinence. By and large, the tests of within-person changes are encouraging in that they suggest that the hoped-for cognitive and emotional changes are indeed taking place as nondaily smokers engage with the SiS2 app. Indeed, interactions with other mHealth smoking cessation technologies have produced similar results [68], although with somewhat weaker effects. Further testing of the SiS2 app in a randomized design is warranted to test the degree to which these changes are attributable to SiS2 app use.

The self-reported 30-day point prevalence abstinence rates we observed are very promising (eg, 56% abstinent at the 6-month follow-up) and certainly exceed expectations for this type of technology. Expected smoking abstinence rates (for daily smokers) are 20%-25% (treatment) versus 14%-16% (control) for SMS text messaging [69,70], and 28% (treatment) versus

21% (control) for an app [62]. However, it must be kept in mind that our study was a single-group longitudinal study, with a strong potential for a response bias, given that there was no blinding and that participants were onboarded by a motivated study staff member. Nevertheless, these high self-reported 30-day abstinence rates indicate that the SiS2 app merits study in a randomized trial.

Limitations

This was a single-group longitudinal study without biochemical verification of smoking status. There is the potential for a response bias in reporting smoking status and potentially regarding our other self-report measures as well. We reminded participants during all stages of the study (ie, during screening, during enrollment, and during survey taking) that their honest reporting was of critical importance to us, even and especially if they did not like something about the app. Note also that our primary feasibility outcome indicator, actual app use, was not subject to self-report biases, as it was automatically recorded as it occurred. Second, our study approach used an interactive onboarding procedure, done via phone, where a staff member was guiding participants through downloading, installing, and using the app. Thus, our results may not generalize to referral-only situations, where smokers find the app on their own or are merely referred to it. We used the onboarding call as it builds on best practices for warm linkages to community resources, similar to a warm handoff for linking hospitalized smokers to quitlines [71], which may be particularly helpful for stigmatized and underserved populations [72]. In addition, our eligibility criteria included willingness to make a quit attempt as part of the study. Motivation is an essential factor in smoking cessation. By restricting study participation to those who are willing to make a quit attempt, this study likely restricted participation to those with a relatively high level of motivation to quit smoking. This focus reflects real-world settings, as individuals do not access smoking cessation programs unless they are motivated. Indeed, national quitlines require that smokers be motivated to quit to use their services. With regard to apps, those not willing to make a quit attempt are unlikely to go to the app store to find a smoking cessation app. The SiS app is designed for those smokers who want to quit smoking and who want to use an app to help them do so. Third, our sample was predominantly White (75/100, 75%), unlike our previous sample in study 1, where Black nondaily smokers were the largest racial group (43%) [24]. Future analyses will examine in depth whether any demographic or clinical characteristics predicted more or less engagement with the SiS2 app.

Conclusions

The SiS2 app was feasible, acceptable, showed promising changes in constructs relevant to smoking cessation, and had high self-reported quit rates by nondaily smokers. The SiS2 app warrants testing in a randomized controlled trial.

Acknowledgments

This research was supported by a grant from the American Cancer Society (grant RSG CPPB-130323) to BBH.

Conflicts of Interest

None declared.

References

1. Creamer MR, Wang TW, Babb S, Cullen KA, Day H, Willis G, et al. Tobacco product use and cessation indicators among adults - United States, 2018. *MMWR Morb Mortal Wkly Rep* 2019 Nov 15;68(45):1013-1019 [FREE Full text] [doi: [10.15585/mmwr.mm6845a2](https://doi.org/10.15585/mmwr.mm6845a2)] [Medline: [31725711](https://pubmed.ncbi.nlm.nih.gov/31725711/)]
2. Centers for Disease Control/Prevention (CDC). Cigarette smoking among adults and trends in smoking cessation - United States, 2008. *MMWR Morb Mortal Wkly Rep* 2009 Nov 13;58(44):1227-1232 [FREE Full text] [Medline: [19910909](https://pubmed.ncbi.nlm.nih.gov/19910909/)]
3. National Cancer Institute. What's in a name? Examination of light and intermittent smokers (NIH Publication). US Department of Health and Human Services. 2007. URL: https://scholar.google.com/scholar?cluster=13594314118933346903&hl=en&as_sdt=40005&scioldt=0,10 [accessed 2021-09-22]
4. Gilpin E, Cavin SW, Pierce J. Adult smokers who do not smoke daily. *Addiction* 1997 Apr;92(4):473-480. [doi: [10.1111/j.1360-0443.1997.tb03379.x](https://doi.org/10.1111/j.1360-0443.1997.tb03379.x)] [Medline: [9177069](https://pubmed.ncbi.nlm.nih.gov/9177069/)]
5. Tong EK, Ong MK, Vittinghoff E, Pérez-Stable EJ. Nondaily smokers should be asked and advised to quit. *Am J Prev Med* 2006 Jan;30(1):23-30. [doi: [10.1016/j.amepre.2005.08.048](https://doi.org/10.1016/j.amepre.2005.08.048)] [Medline: [16414420](https://pubmed.ncbi.nlm.nih.gov/16414420/)]
6. Shiffman S, Dunbar MS, Scholl SM, Tindle HA. Smoking motives of daily and non-daily smokers: a profile analysis. *Drug Alcohol Depend* 2012 Dec 01;126(3):362-368 [FREE Full text] [doi: [10.1016/j.drugalcdep.2012.05.037](https://doi.org/10.1016/j.drugalcdep.2012.05.037)] [Medline: [22784601](https://pubmed.ncbi.nlm.nih.gov/22784601/)]
7. Shiffman S, Dunbar MS, Li X, Scholl SM, Tindle HA, Anderson SJ, et al. Smoking patterns and stimulus control in intermittent and daily smokers. *PLoS One* 2014;9(3):e89911 [FREE Full text] [doi: [10.1371/journal.pone.0089911](https://doi.org/10.1371/journal.pone.0089911)] [Medline: [24599056](https://pubmed.ncbi.nlm.nih.gov/24599056/)]
8. Inoue-Choi M, McNeel TS, Hartge P, Caporaso NE, Graubard BI, Freedman ND. Non-daily cigarette smokers: mortality risks in the U.S. *Am J Prev Med* 2019 Jan;56(1):27-37 [FREE Full text] [doi: [10.1016/j.amepre.2018.06.025](https://doi.org/10.1016/j.amepre.2018.06.025)] [Medline: [30454906](https://pubmed.ncbi.nlm.nih.gov/30454906/)]
9. Inoue-Choi M, Christensen CH, Rostron BL, Cosgrove CM, Reyes-Guzman C, Apelberg B, et al. Dose-response association of low-intensity and nondaily smoking with mortality in the United States. *JAMA Netw Open* 2020 Jun 03;3(6):e206436 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.6436](https://doi.org/10.1001/jamanetworkopen.2020.6436)] [Medline: [32492162](https://pubmed.ncbi.nlm.nih.gov/32492162/)]
10. Trinidad D, Pérez-Stable EJ, Emery S, White M, Grana R, Messer K. Intermittent and light daily smoking across racial/ethnic groups in the United States. *Nicotine Tob Res* 2009 Feb;11(2):203-210 [FREE Full text] [doi: [10.1093/ntr/ntn018](https://doi.org/10.1093/ntr/ntn018)] [Medline: [19246433](https://pubmed.ncbi.nlm.nih.gov/19246433/)]
11. Reyes-Guzman CM, Pfeiffer RM, Lubin J, Freedman ND, Cleary SD, Levine PH, et al. Determinants of light and intermittent smoking in the united states: results from three pooled national health surveys. *Cancer Epidemiol Biomarkers Prev* 2016 Oct 19;26(2):228-239. [doi: [10.1158/1055-9965.epi-16-0028](https://doi.org/10.1158/1055-9965.epi-16-0028)]
12. Blanco L, Garcia R, Pérez-Stable EJ, White MM, Messer K, Pierce JP, et al. National trends in smoking behaviors among mexican, puerto rican, and cuban men and women in the united states. *Am J Public Health* 2014 May;104(5):896-903. [doi: [10.2105/ajph.2013.301844](https://doi.org/10.2105/ajph.2013.301844)]
13. Weinberger AH, Streck JM, Pacek LR, Goodwin RD. Nondaily cigarette smoking is increasing among people with common mental health and substance use problems in the united states: data from representative samples of US adults, 2005-2014. *J Clin Psychiatry* 2018 Aug 14;79(5):m11945 [FREE Full text] [doi: [10.4088/JCP.17m11945](https://doi.org/10.4088/JCP.17m11945)] [Medline: [30153404](https://pubmed.ncbi.nlm.nih.gov/30153404/)]
14. Shiffman S, Scholl S, Mao J, Ferguson S, Hedeker D, Primack B, et al. Using nicotine gum to assist nondaily smokers in quitting: a randomized clinical trial. *Nicotine Tob Res* 2020 Mar 16;22(3):390-397 [FREE Full text] [doi: [10.1093/ntr/ntz090](https://doi.org/10.1093/ntr/ntz090)] [Medline: [31125988](https://pubmed.ncbi.nlm.nih.gov/31125988/)]
15. Nollen NL, Cox LS, Mayo MS, Ellerbeck EF, Ahluwalia JS. Counseling alone or in combination with nicotine replacement therapy for treatment of black non-daily smokers: a randomized trial. *Addiction* 2020 Aug 06;115(8):1547-1560. [doi: [10.1111/add.14948](https://doi.org/10.1111/add.14948)] [Medline: [31899564](https://pubmed.ncbi.nlm.nih.gov/31899564/)]
16. Berg CJ, Sutfin EL, Mendel J, Ahluwalia JS. Use of and interest in smoking cessation strategies among daily and nondaily college student smokers. *J Am Coll Health* 2012 Apr;60(3):194-202 [FREE Full text] [doi: [10.1080/07448481.2011.586388](https://doi.org/10.1080/07448481.2011.586388)] [Medline: [22420696](https://pubmed.ncbi.nlm.nih.gov/22420696/)]
17. Hoepfner B, Hoepfner S, Carlon H, Abry A, Darville A, Rohsenow D. Preparing for the quit day: comparing beliefs of nondaily versus daily young adult smokers as they prepare for a quit attempt. *Nicotine Tob Res* 2021 May 24;23(6):1038-1046 [FREE Full text] [doi: [10.1093/ntr/ntaa166](https://doi.org/10.1093/ntr/ntaa166)] [Medline: [32882037](https://pubmed.ncbi.nlm.nih.gov/32882037/)]
18. Hoepfner BB, Hoepfner SS, Seaboyer L, Schick MR, Wu GW, Bergman BG, et al. How smart are smartphone apps for smoking cessation? A content analysis. *Nicotine Tob Res* 2016 May;18(5):1025-1031 [FREE Full text] [doi: [10.1093/ntr/ntv117](https://doi.org/10.1093/ntr/ntv117)] [Medline: [26045249](https://pubmed.ncbi.nlm.nih.gov/26045249/)]
19. Abroms LC, Westmaas JL, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. *Am J Prev Med* 2013 Dec;45(6):732-736 [FREE Full text] [doi: [10.1016/j.amepre.2013.07.008](https://doi.org/10.1016/j.amepre.2013.07.008)] [Medline: [24237915](https://pubmed.ncbi.nlm.nih.gov/24237915/)]

20. Fiore MC, Baker TB. Clinical practice. Treating smokers in the health care setting. *N Engl J Med* 2011 Sep 29;365(13):1222-1231 [[FREE Full text](#)] [doi: [10.1056/NEJMcp1101512](https://doi.org/10.1056/NEJMcp1101512)] [Medline: [21991895](https://pubmed.ncbi.nlm.nih.gov/21991895/)]
21. Hoeppepner BB, Hoeppepner SS, Kelly L, Schick M, Kelly JF. Smiling instead of smoking: development of a positive psychology smoking cessation smartphone app for non-daily smokers. *Int J Behav Med* 2017 Oct 14;24(5):683-693. [doi: [10.1007/s12529-017-9640-9](https://doi.org/10.1007/s12529-017-9640-9)] [Medline: [28197846](https://pubmed.ncbi.nlm.nih.gov/28197846/)]
22. Kahler CW, Spillane NS, Day A, Clerkin EM, Parks A, Leventhal AM, et al. Positive psychotherapy for smoking cessation: treatment development, feasibility and preliminary results. *J Posit Psychol* 2014 Jan 01;9(1):19-29 [[FREE Full text](#)] [doi: [10.1080/17439760.2013.826716](https://doi.org/10.1080/17439760.2013.826716)] [Medline: [24683417](https://pubmed.ncbi.nlm.nih.gov/24683417/)]
23. Kahler CW, Spillane NS, Day AM, Cioe PA, Parks A, Leventhal AM, et al. Positive psychotherapy for smoking cessation: a pilot randomized controlled trial. *Nicotine Tob Res* 2015 Nov 02;17(11):1385-1392 [[FREE Full text](#)] [doi: [10.1093/ntr/ntv011](https://doi.org/10.1093/ntr/ntv011)] [Medline: [25646352](https://pubmed.ncbi.nlm.nih.gov/25646352/)]
24. Hoeppepner BB, Hoeppepner SS, Carlon HA, Perez GK, Helmuth E, Kahler CW, et al. Leveraging positive psychology to support smoking cessation in nondaily smokers using a smartphone app: feasibility and acceptability study. *JMIR Mhealth Uhealth* 2019 Jul 03;7(7):e13436 [[FREE Full text](#)] [doi: [10.2196/13436](https://doi.org/10.2196/13436)] [Medline: [31271147](https://pubmed.ncbi.nlm.nih.gov/31271147/)]
25. Locke EA. Social foundations of thought and action: a social-cognitive view. *Acad Manag Rev* 1987 Jan;12(1):169-171. [doi: [10.5465/amr.1987.4306538](https://doi.org/10.5465/amr.1987.4306538)]
26. Moskowitz JT, Carrico AW, Duncan LG, Cohn MA, Cheung EO, Batchelder A, et al. Randomized controlled trial of a positive affect intervention for people newly diagnosed with HIV. *J Consult Clin Psychol* 2017 May;85(5):409-423 [[FREE Full text](#)] [doi: [10.1037/ccp0000188](https://doi.org/10.1037/ccp0000188)] [Medline: [28333512](https://pubmed.ncbi.nlm.nih.gov/28333512/)]
27. Huffman JC, Mastromauro CA, Boehm JK, Seabrook R, Fricchione GL, Denninger JW, et al. Development of a positive psychology intervention for patients with acute cardiovascular disease. *Heart Int* 2011 Sep 29;6(2):e14 [[FREE Full text](#)] [doi: [10.4081/hi.2011.e14](https://doi.org/10.4081/hi.2011.e14)] [Medline: [23825741](https://pubmed.ncbi.nlm.nih.gov/23825741/)]
28. Ogedegbe GO, Boutin-Foster C, Wells MT, Allegrante JP, Isen AM, Jobe JB, et al. A randomized controlled trial of positive-affect intervention and medication adherence in hypertensive African Americans. *Arch Intern Med* 2012 Feb 27;172(4):322-326 [[FREE Full text](#)] [doi: [10.1001/archinternmed.2011.1307](https://doi.org/10.1001/archinternmed.2011.1307)] [Medline: [22269592](https://pubmed.ncbi.nlm.nih.gov/22269592/)]
29. Carrico AW, Gómez W, Siever MD, Discepolo MV, Dilworth SE, Moskowitz JT. Pilot randomized controlled trial of an integrative intervention with methamphetamine-using men who have sex with men. *Arch Sex Behav* 2015 Oct 30;44(7):1861-1867 [[FREE Full text](#)] [doi: [10.1007/s10508-015-0505-5](https://doi.org/10.1007/s10508-015-0505-5)] [Medline: [26123068](https://pubmed.ncbi.nlm.nih.gov/26123068/)]
30. Hoeppepner BB, Kahler CW, Gwaltney CJ. Relationship between momentary affect states and self-efficacy in adolescent smokers. *Health Psychol* 2014 Dec;33(12):1507-1517 [[FREE Full text](#)] [doi: [10.1037/hea0000075](https://doi.org/10.1037/hea0000075)] [Medline: [25020151](https://pubmed.ncbi.nlm.nih.gov/25020151/)]
31. Shiffman S, Dunbar MS, Kirchner TR, Li X, Tindle HA, Anderson SJ, et al. Cue reactivity in non-daily smokers: effects on craving and on smoking behavior. *Psychopharmacology (Berl)* 2013 Mar;226(2):321-333 [[FREE Full text](#)] [doi: [10.1007/s00213-012-2909-4](https://doi.org/10.1007/s00213-012-2909-4)] [Medline: [23142992](https://pubmed.ncbi.nlm.nih.gov/23142992/)]
32. Rabois D, Haaga DA. The influence of cognitive coping and mood on smokers' self-efficacy and temptation. *Addict Behav* 2003 Apr;28(3):561-573. [doi: [10.1016/s0306-4603\(01\)00249-0](https://doi.org/10.1016/s0306-4603(01)00249-0)] [Medline: [12628627](https://pubmed.ncbi.nlm.nih.gov/12628627/)]
33. Das E, Vonkeman C, Hartmann T. Mood as a resource in dealing with health recommendations: how mood affects information processing and acceptance of quit-smoking messages. *Psychol Health* 2012 Jan;27(1):116-127. [doi: [10.1080/08870446.2011.569888](https://doi.org/10.1080/08870446.2011.569888)] [Medline: [21678163](https://pubmed.ncbi.nlm.nih.gov/21678163/)]
34. 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/)]
35. Bandura A. Social cognitive theory of self-regulation. *Organ Behav Hum Decis Process* 1991 Dec;50(2):248-287. [doi: [10.1016/0749-5978\(91\)90022-L](https://doi.org/10.1016/0749-5978(91)90022-L)]
36. Ajzen I, Madden TJ. Prediction of goal-directed behavior: attitudes, intentions, and perceived behavioral control. *J Exp Soc Psychol* 1986 Sep;22(5):453-474. [doi: [10.1016/0022-1031\(86\)90045-4](https://doi.org/10.1016/0022-1031(86)90045-4)]
37. 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)]
38. Marlatt GA, Donovan DM. *Maintenance Strategies in the Treatment of Addictive Behaviors*. New York, NY, United States: The Guilford Press; 2005:1-416.
39. 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/)]
40. Krentzman AR, Mannella KA, Hassett AL, Barnett NP, Cranford JA, Brower KJ, et al. Feasibility, acceptability, and impact of a web-based gratitude exercise among individuals in outpatient treatment for alcohol use disorder. *J Posit Psychol* 2015 Feb 25;10(6):477-488 [[FREE Full text](#)] [doi: [10.1080/17439760.2015.1015158](https://doi.org/10.1080/17439760.2015.1015158)] [Medline: [27076837](https://pubmed.ncbi.nlm.nih.gov/27076837/)]
41. Hoeppepner BB, Schick MR, Carlon H, Hoeppepner SS. Do self-administered positive psychology exercises work in persons in recovery from problematic substance use? An online randomized survey. *J Subst Abuse Treat* 2019 Apr;99:16-23. [doi: [10.1016/j.jsat.2019.01.006](https://doi.org/10.1016/j.jsat.2019.01.006)] [Medline: [30797389](https://pubmed.ncbi.nlm.nih.gov/30797389/)]
42. Lyubomirsky S, King L, Diener E. The benefits of frequent positive affect: does happiness lead to success? *Psychol Bull* 2005 Nov;131(6):803-855. [doi: [10.1037/0033-2909.131.6.803](https://doi.org/10.1037/0033-2909.131.6.803)] [Medline: [16351326](https://pubmed.ncbi.nlm.nih.gov/16351326/)]

43. Otake K, Shimai S, Tanaka-Matsumi J, Otsui K, Fredrickson BL. Happy people become happier through kindness: a counting kindnesses intervention. *J Happiness Stud* 2006 Sep;7(3):361-375 [FREE Full text] [doi: [10.1007/s10902-005-3650-z](https://doi.org/10.1007/s10902-005-3650-z)] [Medline: [17356687](https://pubmed.ncbi.nlm.nih.gov/17356687/)]
44. Jose PE, Lim BT, Bryant FB. Does savoring increase happiness? A daily diary study. *J Posit Psychol* 2012 May;7(3):176-187. [doi: [10.1080/17439760.2012.671345](https://doi.org/10.1080/17439760.2012.671345)]
45. Fiore MC, Jaén CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, et al. Treating tobacco use and dependence: 2008 update. U.S. Department of Health and Human Services. 2008. URL: <http://www.tobaccoprogram.org/clientuploads/documents/Consumer%20Materials/Clinicians%20Systems%20Mat/2008-Guidelines.pdf> [accessed 2021-09-22]
46. McClure JB, Hartzler AL, Catz SL. Design considerations for smoking cessation apps: feedback from nicotine dependence treatment providers and smokers. *JMIR Mhealth Uhealth* 2016 Feb 12;4(1):e17 [FREE Full text] [doi: [10.2196/mhealth.5181](https://doi.org/10.2196/mhealth.5181)] [Medline: [26872940](https://pubmed.ncbi.nlm.nih.gov/26872940/)]
47. Brooke J. SUS: A 'quick and dirty' usability. In: *Usability Evaluation In Industry*. Boca Raton, Florida, United States: CRC Press; 1996.
48. Sauro J. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. California, US: Createspace Independent Pub; Apr 20, 2011.
49. Etter J, Bergman M, Humair J, Perneger T. Development and validation of a scale measuring self-efficacy of current and former smokers. *Addiction* 2000 Jun;95(6):901-913. [doi: [10.1046/j.1360-0443.2000.9569017.x](https://doi.org/10.1046/j.1360-0443.2000.9569017.x)] [Medline: [10946439](https://pubmed.ncbi.nlm.nih.gov/10946439/)]
50. Cox LS, Tiffany ST, Christen AG. Evaluation of the brief questionnaire of smoking urges (qsu-brief) in laboratory and clinical settings. *Nicotine Tob Res* 2001 Feb;3(1):7-16. [doi: [10.1080/14622200020032051](https://doi.org/10.1080/14622200020032051)] [Medline: [11260806](https://pubmed.ncbi.nlm.nih.gov/11260806/)]
51. Etter J, Humair J, Bergman M, Perneger T. Development and validation of the attitudes towards smoking scale (ATS-18). *Addiction* 2000 Apr;95(4):613-625. [doi: [10.1046/j.1360-0443.2000.95461312.x](https://doi.org/10.1046/j.1360-0443.2000.95461312.x)] [Medline: [10829336](https://pubmed.ncbi.nlm.nih.gov/10829336/)]
52. Ward RM, Velicer WF, Rossi JS, Fava JL, Prochaska JO. Factorial invariance and internal consistency for the decisional balance inventory--short form. *Addict Behav* 2004 Jul;29(5):953-958. [doi: [10.1016/j.addbeh.2004.02.042](https://doi.org/10.1016/j.addbeh.2004.02.042)] [Medline: [15219341](https://pubmed.ncbi.nlm.nih.gov/15219341/)]
53. Kahler CW, Lachance HR, Strong DR, Ramsey SE, Monti PM, Brown RA. The commitment to quitting smoking scale: initial validation in a smoking cessation trial for heavy social drinkers. *Addict Behav* 2007 Oct;32(10):2420-2424 [FREE Full text] [doi: [10.1016/j.addbeh.2007.04.002](https://doi.org/10.1016/j.addbeh.2007.04.002)] [Medline: [17478057](https://pubmed.ncbi.nlm.nih.gov/17478057/)]
54. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the panas scales. *J Pers Soc Psychol* 1988;54(6):1063-1070. [doi: [10.1037/0022-3514.54.6.1063](https://doi.org/10.1037/0022-3514.54.6.1063)]
55. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. *J Pers Assess* 1985 Feb;49(1):71-75. [doi: [10.1207/s15327752jpa4901_13](https://doi.org/10.1207/s15327752jpa4901_13)] [Medline: [16367493](https://pubmed.ncbi.nlm.nih.gov/16367493/)]
56. Lyubomirsky S, Lepper HS. A measure of subjective happiness: preliminary reliability and construct validation. *Soc Indic Res* 1999 Feb;46(2):137-155. [doi: [10.1023/A:1006824100041](https://doi.org/10.1023/A:1006824100041)]
57. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. *Nicotine Tob Res* 2002 May 1;4(2):149-159. [doi: [10.1080/14622200210123581](https://doi.org/10.1080/14622200210123581)] [Medline: [12028847](https://pubmed.ncbi.nlm.nih.gov/12028847/)]
58. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and anovas. *Front Psychol* 2013 Nov 26;4:863 [FREE Full text] [doi: [10.3389/fpsyg.2013.00863](https://doi.org/10.3389/fpsyg.2013.00863)] [Medline: [24324449](https://pubmed.ncbi.nlm.nih.gov/24324449/)]
59. Shiffman S, Tindle H, Li X, Scholl S, Dunbar M, Mitchell-Miland C. Characteristics and smoking patterns of intermittent smokers. *Exp Clin Psychopharmacol* 2012 Aug;20(4):264-277 [FREE Full text] [doi: [10.1037/a0027546](https://doi.org/10.1037/a0027546)] [Medline: [22390208](https://pubmed.ncbi.nlm.nih.gov/22390208/)]
60. 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/)]
61. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. *Psychiatr Serv* 2017 May 01;68(5):427-429 [FREE Full text] [doi: [10.1176/appi.ps.201600541](https://doi.org/10.1176/appi.ps.201600541)] [Medline: [28412890](https://pubmed.ncbi.nlm.nih.gov/28412890/)]
62. Bricker JB, Watson NL, Mull KE, Sullivan BM, Heffner JL. Efficacy of smartphone applications for smoking cessation: a randomized clinical trial. *JAMA Intern Med* 2020 Nov 01;180(11):1472-1480. [doi: [10.1001/jamainternmed.2020.4055](https://doi.org/10.1001/jamainternmed.2020.4055)] [Medline: [32955554](https://pubmed.ncbi.nlm.nih.gov/32955554/)]
63. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. a U.S. public health service report. *Am J Prev Med* 2008 Aug 01;35(2):158-176 [FREE Full text] [doi: [10.1016/j.amepre.2008.04.009](https://doi.org/10.1016/j.amepre.2008.04.009)] [Medline: [18617085](https://pubmed.ncbi.nlm.nih.gov/18617085/)]
64. 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/)]
65. Vilardaga R, Rizo J, Kientz JA, McDonnell MG, Ries RK, Sobel K. User experience evaluation of a smoking cessation app in people with serious mental illness. *Nicotine Tob Res* 2016 May;18(5):1032-1038 [FREE Full text] [doi: [10.1093/ntr/ntv256](https://doi.org/10.1093/ntr/ntv256)] [Medline: [26581430](https://pubmed.ncbi.nlm.nih.gov/26581430/)]

66. Vilardaga R, Rizo J, Palenski PE, Mannelli P, Oliver JA, Mcclernon FJ. Pilot randomized controlled trial of a novel smoking cessation app designed for individuals with co-occurring tobacco use disorder and serious mental illness. *Nicotine Tob Res* 2020 Aug 24;22(9):1533-1542 [FREE Full text] [doi: [10.1093/ntr/ntz202](https://doi.org/10.1093/ntr/ntz202)] [Medline: [31667501](https://pubmed.ncbi.nlm.nih.gov/31667501/)]
67. Curry S, Marlatt GA, Gordon JR. Abstinence violation effect: validation of an attributional construct with smoking cessation. *J Consult Clin Psychol* 1987 Apr;55(2):145-149. [doi: [10.1037/0022-006x.55.2.145](https://doi.org/10.1037/0022-006x.55.2.145)]
68. Hoeppepner BB, Hoeppepner SS, Schick MR, Milligan CM, Helmuth E, Bergman BG, et al. Using the text-messaging program smokefreetxt to support smoking cessation for nondaily smokers. *Subst Use Misuse* 2019 Apr 19;54(8):1260-1271. [doi: [10.1080/10826084.2018.1552300](https://doi.org/10.1080/10826084.2018.1552300)] [Medline: [30999794](https://pubmed.ncbi.nlm.nih.gov/30999794/)]
69. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. *Lancet* 2011 Jul;378(9785):49-55. [doi: [10.1016/S0140-6736\(11\)60701-0](https://doi.org/10.1016/S0140-6736(11)60701-0)]
70. Abroms LC, Boal AL, Simmens SJ, Mendel JA, Windsor RA. A randomized trial of text2quit: a text messaging program for smoking cessation. *Am J Prev Med* 2014 Sep;47(3):242-250 [FREE Full text] [doi: [10.1016/j.amepre.2014.04.010](https://doi.org/10.1016/j.amepre.2014.04.010)] [Medline: [24913220](https://pubmed.ncbi.nlm.nih.gov/24913220/)]
71. Richter KP, Faseru B, Shireman TI, Mussulman LM, Nazir N, Bush T, et al. Warm handoff versus fax referral for linking hospitalized smokers to quitlines. *Am J Prev Med* 2016 Oct;51(4):587-596 [FREE Full text] [doi: [10.1016/j.amepre.2016.04.006](https://doi.org/10.1016/j.amepre.2016.04.006)] [Medline: [27647059](https://pubmed.ncbi.nlm.nih.gov/27647059/)]
72. Mussulman LM, Faseru B, Fitzgerald S, Nazir N, Patel V, Richter KP. A randomized, controlled pilot study of warm handoff versus fax referral for hospital-initiated smoking cessation among people living with hiv/aids. *Addict Behav* 2018 Mar;78:205-208 [FREE Full text] [doi: [10.1016/j.addbeh.2017.11.035](https://doi.org/10.1016/j.addbeh.2017.11.035)] [Medline: [29216569](https://pubmed.ncbi.nlm.nih.gov/29216569/)]

Abbreviations

mHealth: mobile health
NCI: National Cancer Institute
NRT: nicotine replacement therapy
REDCap: Research Electronic Data Capture
SiS: Smiling Instead of Smoking
SUS: System Usability Scale

Edited by G Eysenbach; submitted 20.04.21; peer-reviewed by N Machado, N Krishnan; comments to author 30.07.21; revised version received 12.08.21; accepted 24.08.21; published 17.11.21.

Please cite as:

Hoeppepner BB, Siegel KR, Carlon HA, Kahler CW, Park ER, Hoeppepner SS
A Smoking Cessation App for Nondaily Smokers (Version 2 of the Smiling Instead of Smoking App): Acceptability and Feasibility Study
JMIR Form Res 2021;5(11):e29760
URL: <https://formative.jmir.org/2021/11/e29760>
doi: [10.2196/29760](https://doi.org/10.2196/29760)
PMID: [34787577](https://pubmed.ncbi.nlm.nih.gov/34787577/)

©Bettina B Hoeppepner, Kaitlyn R Siegel, Hannah A Carlon, Christopher W Kahler, Elyse R Park, Susanne S Hoeppepner. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 17.11.2021. 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

Teleassessments for Enrollment of Adults With Physical or Mobility Disability in a Home-Based Exercise Trial in Response to COVID-19: Usability Study

Jereme Wilroy^{1*}, PhD; Byron Lai^{1*}, PhD; Madison Currie¹, BSc; Hui-Ju Young¹, PhD; Mohanraj Thirumalai¹, PhD; Tapan Mehta¹, PhD; John Giannone¹, MSc; James Rimmer¹, PhD

University of Alabama at Birmingham, Birmingham, AL, United States

*these authors contributed equally

Corresponding Author:

Jereme Wilroy, PhD

University of Alabama at Birmingham

1717 6th Ave S

Birmingham, AL, 35233

United States

Phone: 1 2059344508

Email: jdwilroy@uab.edu

Abstract

Background: The Scale-Up Project Evaluating Responsiveness to Home Exercise And Lifestyle Tele-Health (SUPER-HEALTH) initiative is a large randomized controlled study that aims to overcome logistical barriers to exercise via telehealth for people with physical disabilities. However, at the start of the COVID-19 pandemic, enrollment was halted due to limited operations at the testing site, which included no onsite visits that involved participant data collection. In response to the limited operations, a modified data collection protocol was developed for virtual enrollment of study participants.

Objective: This paper presents feasibility data on using teleassessments to enroll people with mobility impairment into a home-based exercise trial.

Methods: The modified protocol replaced onsite enrollment and data collection visits with teleassessments using a computer tablet and testing equipment that was shipped to the participants' home address prior to the synchronous teleassessments conducted by an exercise physiologist through Zoom. The participants were mailed a teleassessment toolkit that included a digital blood pressure cuff, spirometer, hand dynamometer, mini disc cone, and measuring tape (to complete standardized testing). The teleassessment measures included resting blood pressure and heart rate, forced vital capacity, grip strength, Five Times Sit to Stand, and Timed Up and Go. Feasibility metrics included technological effectiveness, efficiency, and safety. The technological effectiveness of the telehealth assessment was determined by the percentage of sessions completed without technical issues with $\geq 90\%$ criteria set a priori. Efficiency was measured by a session duration of ≤ 2 hours. Safety was measured by the number of adverse events related to the teleassessments reported.

Results: Data from 36 participants were included in this feasibility study, and 34 (94%) participants completed all teleassessments without technical issues. For efficiency, the teleassessment sessions were completed in a mean time of 65 minutes and a maximum session length of 110 minutes. There were no adverse events reported to indicate concerns with the safety of teleassessments.

Conclusions: The modified teleassessment protocol, in response to COVID-19 restrictions, may be a feasible process for enrolling adults with mobility impairment into a home exercise trial who otherwise would have not been able to participate.

Trial Registration: ClinicalTrials.gov NCT03024320; <https://clinicaltrials.gov/ct2/show/NCT03024320>

(*JMIR Form Res* 2021;5(11):e29799) doi:[10.2196/29799](https://doi.org/10.2196/29799)

KEYWORDS

telehealth; disability; COVID-19; exercise; assessments; feasibility; mHealth; teleassessment; mobility impairment; home exercise; participation; physical disabilities

Introduction

People with physical disabilities are at greater risk for primary and secondary health conditions compared with people without a disability [1-3]. These include primary health conditions, such as heart disease and diabetes, and secondary conditions, including pain, depression, sleep disturbance, spasticity, and many others [4,5]. Despite physical activity becoming an imperative public health priority for all demographics, there are still many existing barriers to physical activity. A few of these barriers include accessibility of facilities, opportunity for activity, and aesthetic or environmental attributes [6]. In order to overcome these barriers, people with physical disabilities must be given more options for home-based activity, thus eliminating many well-known community-based barriers [7,8].

The SUPER-HEALTH project, which stands for Scale-Up Project Evaluating Responsiveness to Home Exercise And Lifestyle Tele-Health, is an ongoing 48-week exercise trial assessing the utilization of a movement-to-music (M2M) intervention remotely delivered to various disability groups [8]. SUPER-HEALTH recruits individuals with a mobility impairment, defined as the inability to walk or difficulty walking as a means to exercise, and includes people with spinal cord injury (SCI), spina bifida (SB), multiple sclerosis (MS), stroke, limb loss, and other disabilities and chronic health conditions. SUPER-HEALTH was designed to deliver an exercise training system in the convenience of the home in order to avoid barriers associated with transportation for onsite exercising at a facility. However, project enrollment was halted by the University of Alabama at Birmingham due to COVID-19 restrictions involving onsite data collection. To overcome this issue, the project team developed teleassessment protocols that enabled the trial to continue enrolling eligible participants across the southeastern United States [9]. The purpose of this paper is to describe the feasibility of using teleassessments for enrolling people with physical disabilities into a home-based exercise trial by evaluating three feasibility components, namely technological effectiveness (ie, successful implementation without technical error), efficiency, and safety.

Textbox 1. Inclusion and exclusion criteria.

<p>Inclusion criteria</p> <ul style="list-style-type: none">• Self-report of a physical disability or mobility impairment• 18 to 74 years of age• Not currently enrolled in a structured exercise program over the past 6 months• Have the ability to use upper, lower, or both sets of extremities to exercise• Have the ability to converse and read English• Agree to receiving emails to complete consent, surveys, and teleassessments <p>Exclusion criteria</p> <ul style="list-style-type: none">• Medically unstable to perform home exercise as determined by their physician• Cognitive impairment that may preclude self-directed daily activities• No wireless internet in home (can include hotspot with unlimited data)
--

Methods

Description of SUPER-HEALTH

SUPER-HEALTH is testing a 48-week, home-based exercise training system among a large sample (N=648) of people with mobility impairments by comparing 3 groups: (1) attention control group (AC) receiving health promotion articles; (2) exercise group (M2M) receiving exercise videos and health promotion articles; and (3) exercise group receiving exercise videos, health promotion articles, and social networking features (M2M^{plus}). Participants in each group received a Samsung Tab A 10.1 tablet, which was inserted into a protective case and customized with features pertaining to the study. These features included downloads of the study app [7], the Fitbit app with the Fitbit Charge 4 synced and connected by Bluetooth, and the Zoom app for communication. An email address was provided to the participant along with a Gmail widget configured on the tablet to communicate with research staff. Additionally, the exercise groups also received wrist weights.

This protocol was approved by the university's Institutional Review Board (IRB) and is registered with ClinicalTrials.gov (NCT03024320) as a phase III clinical trial. We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during the course of this research.

Participants

Participants who were part of the SUPER-HEALTH study and who completed teleassessment were analyzed as part of this feasibility study. Textbox 1 provides lists of inclusion and exclusion criteria, which are the same as the main trial with the additional criteria concerning the completion of teleassessments. A convenience sample size of 36 was chosen to satisfy technological usability [10] and feasibility study recommendations [11]. In order to recruit participants for the study, a list of patient names and addresses from a health informatics database was obtained and study brochures were mailed out. The participants enrolled by calling, providing contact information through the study website, or contacting the study email.

COVID-19 Modifications

In order to avoid the need for participants to commute to the laboratory for onsite data collection, our team developed protocols for synchronous remote assessments (teleassessments) and virtual orientation. The new teleassessment protocol was used for the original study from this point forward. In the following section, we will discuss some aspects of the modified protocol.

Consent and Survey Completion

As with the original study protocol, consent and survey data collection involved emailing links to participants, connecting them to an electronic form via REDCap (electronic data capture system). Once the participants signed and submitted the electronic consent, they would automatically receive surveys from REDCap. The questionnaires included a demographic and health history questionnaire; a medication list; Godin Leisure Time Exercise Questionnaire (GLTEQ) [12]; social cognitive theory questionnaires for exercise self-efficacy [13], barriers to physical activity [14], outcome expectations [15], and social support [16]; a set of Patient Reported Outcomes Measurement Information System (PROMIS) short forms [17] to assess quality of life; and an eHealth Literacy Scale [18]. After all of the

surveys were completed, the research staff packed and shipped the assessment and training equipment to the participants' provided home address, as part of the modified protocol.

Teleassessment Package

Teleassessment equipment mailed to each participant included a digital blood pressure and heart rate monitor (US \$45), hand-held dynamometer (US \$30), spirometer (US \$25), mini disc cone (US \$1), and soft measuring tape (US \$2). All of the equipment was lightweight and ready to use out of the box due to preparation by a research assistant prior to shipping. Teleassessment equipment was packaged along with the Samsung Tab A tablet (US \$200) and Fitbit Charge 4 device (US \$120) (Figure 1). Each participant was informed of equipment shipment date and expected delivery date, and the package was tracked by the research assistant. A brief instruction sheet was included and simply stated to keep the tablet charged, await call from tester to schedule virtual testing appointment, and email that equipment has been received, along with the option to contact study staff with any questions. Once the participant received their equipment, the teleassessor reached out to schedule the participant for their baseline teleassessment session. The participants kept all equipment for the duration of the study and beyond.

Figure 1. Teleassessment equipment, which included a digital blood pressure cuff, a hand dynamometer, a spirometer, a mini disc cone, and a 3-meter measuring tape.



Teleassessment Session

A single teleassessor, an exercise and sport science lab assistant with a master's in exercise physiology, was used to complete the teleassessment session. At the start of the session, the teleassessor instructed the participant to make sure the tablet was on a table propped on case stand and ensured the participant was in full view. They would then go through a script concerning a plan for adverse events (eg, fall) and internet connectivity issues. The session included the following assessments: anthropometrics (2 resting blood pressure readings, 2 resting heart rate readings), grip strength, peak expiratory

volume, Five Times Sit to Stand, and Timed Up and Go (Figure 2). All measures were completed based on the participants' disability and level of function. For example, if someone was unable to walk, they would not complete the Five Times Sit to Stand and Timed Up and Go tests. For exercise groups, the participant's functional group was assigned to ensure they received the appropriate exercise videos, which was (1) able to stand, (2) hemiparesis, or (3) sitting only. This assignment was based on the participant's diagnosis, 14-second cut-point of Timed Up and Go (ie, assigned to seated exercise group if greater than 14 seconds), and participant preference.

Figure 2. An example of a person completing grip strength assessment with teleassessor on videoconference.



Virtual Equipment Orientation

Once the teleassessment was completed, the participant received a virtual equipment orientation via Zoom to notify which group they had been allocated to. They were also mailed a US \$25 Visa gift card via UPS mail delivery. The purpose of this meeting was to inform the participants of which group they had been randomized to and how to use each piece of equipment correctly (ie, Fitbit and tablet) for the purposes of the intervention.

Data Collection for Feasibility

Since the COVID-19 modifications were focused on the application of a technological medium (videoconferencing), we chose to assess the feasibility of the teleassessment process [19] through 3 metrics of technological usability: technological effectiveness, efficiency, and safety [10].

Technological Effectiveness

Effectiveness was defined as the percentage of sessions completed successfully without technical or logistical difficulty. These difficulties could have included any issues with

information communication technology (ie, tablet, internet, videoconference software, or assessment equipment) experienced by the participant or teleassessor. Physiologic testing issues, such as high resting blood pressure or heart rate or an inability to complete an assessment due to mobility level (eg, wheelchair use), were excluded because these cases would have also excluded the participants from completing assessments onsite. We established an a priori criterion for acceptable effectiveness at 90% used previously with this population [20]. A high criterion of 90% was chosen because we wanted the participants to have minimal issues and a high-quality experience during the teleassessments.

Efficiency

Efficiency was defined as the amount of time to complete teleassessments. The timer started once the participant and teleassessor logged on to the videoconferencing session and ended at the point of logging off. The criterion for reasonable efficiency was set at 2 hours or less a priori as this was the time allotted for onsite visits.

Table 1. Demographic data of participants (N=36).

Characteristic	Value
Age (years), mean (SD)	51.08 (16.64)
Gender, n (%)	
Male	11 (31)
Female	25 (69)
Race, n (%)	
African American	21 (58)
White	13 (36)
Other	2 (6)
Disability, n (%)	
Spinal cord injury	5 (14)
Spina bifida	5 (14)
Arthritis	11 (31)
Stroke	3 (8)
Other	12 (33)
Device, n (%)	
Wheelchair only	6 (17)
Cane only	6 (17)
Motorized scooter only	2 (6)
Other	1 (3)
No assistive devices	11 (31)

Feasibility

Technological Effectiveness

Out of the 36 participants in the sample, 34 (94%) completed all teleassessments without issues, which met the criteria for acceptability (>90%) set a priori. This resulted in the successful completion of 188 assessments. The reasons for the 2 participants not completing all teleassessments included:

Safety

Adverse events, defined by the university's IRB as "...any untoward or unfavorable medical occurrence in a human subject," were reported by the teleassessor with a written description of the event. The event was then recorded in a medical oversight form, which was reviewed by the study physician who classified the event as serious or non-serious and whether it was due to study procedures. Any adverse event was then reported to the IRB. Examples of adverse events would include falls or injury during testing.

Results

Demographics

Data from a convenience sample of 36 people were included within this feasibility report. The participants were primarily female and African American (Table 1). These participants were located in southeastern United States.

incomplete spirometry test due to confusion regarding manual peak flow meter and incomplete Timed Up and Go assessment due to technical difficulties with videoconferencing on the tablet.

Efficiency

The mean time for the completion of a teleassessment session for the sample of 36 participants was 65.03 (SD 15.51) minutes, which met our a priori criterion for acceptability (maximum

session length: 110.00 minutes). Therefore, the criterion for completing the sessions in under 2 hours was met by all participants (N=36, 100%) of the sample. The reasons reported for the longest session times included tablet not being charged prior to the visit and internet connection issues.

Safety

There were no adverse events reported from the sample of 36 teleassessment sessions.

Discussion

Principal Considerations

Many trials have been halted due to the restriction of onsite visits by research participants during the COVID-19 pandemic. Given that a significant percentage of the population is either not yet vaccinated or elects not to be vaccinated [21], exercise testing may need to continue to be remotely conducted in the foreseeable future. In addition, other rehabilitation researchers have endorsed the use of functional assessments not requiring costly or complex equipment or specialized training [22]. There is a strong need for researchers to have effective, safe teleassessment protocols that are equivalent to in-person testing.

SUPER-HEALTH is inclusive to varying levels of disability and functionality of all participating users. Thus, the use of telecommunication, teleintervention, and interactive methods were specifically designed to collect the same data virtually as it is carried out onsite. Along with these modifications, the virtual face-to-face interactions between the participants and the study staff suggest that telecommunication is still personable and equally as effective [23]. Prior to COVID-19, most of the enrollment process was completed remotely: phone screen, obtaining medical clearance, electronic consent, and electronic surveys. However, the physiological assessments as well as the equipment orientation were still conducted onsite, which often led to several missed visits and individuals lost to contact. The newly modified protocol utilizing telehealth to conduct these assessments and orientation has enabled the enrollment process to be completed remotely. This eliminated the need for transportation, which is the primary barrier to accessing community exercise programs among patients with physical disability [6]. The teleassessment package was affordable at around US \$100. In addition, the teleassessment session duration was almost half that of the onsite, which, along with fewer missed sessions, saved cost for lab assistant's time. Lastly, the cost for using lab space was substantially more than the smaller tele-suite the lab assistant was able to use for the teleassessment sessions.

These findings are an extension of our work related to modifying specific fitness assessment protocols for people with disabilities to be completed remotely [9]. One thing to note is that elements of the teleassessments, such as laptop camera angle and variability of testing environment and setups, might affect the validity and reliability of the tests. Future research on teleassessments should focus on modifying more testing protocols to ensure the reliability and validity of the data collected. Additionally, future studies should compare data collected onsite with data collected through teleassessments for

any statistically significant differences. A preliminary analysis of the standard deviations between teleassessments and onsite assessments with the current sample did not indicate significant differences; however, a higher-powered validity comparison study and a test-retest study should be completed.

Implications

A challenge often found in exercise trials involving people with disabilities is recruitment. One of the most limiting factors is enrolling large samples in exercise research for this population. A scoping review reported a lack of descriptive details on study participants and noted that many of the published studies have small sample sizes, which were primarily due to individuals being excluded, declining to participate, or dropping out [24]. Barriers reported by patients with physical disability to enrolling in an exercise trial include transportation, scheduling conflicts, secondary health conditions, and difficulty starting a program [25]. The utilization of teleassessments allows researchers to circumnavigate these reported barriers by increasing their ability to enroll more participants. For example, prior to modifying the assessment procedures to enable remote assessment for SUPER-HEALTH, a primary barrier reported by participants was transportation to the research lab. However, the teleassessment protocols have more than doubled the rate of recruitment, and participants are now able to enroll from the convenience of their home. Although 1 onsite laboratory test (ie, submaximal cardiorespiratory endurance) could not be performed remotely and was excluded from the teleassessment protocol, all of the other functional and self-reported measures were successfully conducted remotely. Our current battery of tests provides a starting place for remotely collecting data from the participants. While many studies had to temporarily shut down due to the COVID-19 pandemic, we were able to continue our research by implementing a teleassessment protocol that was found to be effective, efficient, and safe. Future research involving hard-to-reach populations (ie, people with disabilities) should continue to explore the use of teleassessment and tele-exercise in order to reach a broader, more generalizable segment of the population [22].

Limitations

It is important to recognize the limitations of the study. First, the study excluded people who did not have internet access at home, which may have affected the representativeness of the sample. Second, the study included 12 participants who were non-ambulatory (wheelchair users) and were unable to complete two of the functional tests (Five Times Sit to Stand and Timed Up and Go), which reduced the time for completing the teleassessment session. Third, although the sample size meets the recommendations for feasibility studies [10,11], it may not be fully representative of the final sample size of the larger trial, which will be determined at the end of the trial to inform whether data can be merged. Fourth, since the assessments were delivered through telehealth, it is important to note the possible limitations in generalizability due to an inability to account for contextual factors varying among the participants' location of testing [23]. Finally, the results should be interpreted with caution. This is because the study outcomes were chosen based upon arbitrary criteria, and there are no established criterion for

the study outcomes that are considered to be feasible or acceptable.

Conclusions

The use of teleassessments to enroll patients with physical disability into an exercise trial may be technologically effective,

efficient, and safe. Due to the COVID-19 pandemic, many clinical trials have been suspended. The procedures described in this study can be replicated by researchers and health professionals to circumvent the barriers to conducting exercise trials during the pandemic.

Acknowledgments

This research was sponsored by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (5R01HD085186) and a Mentored Career Development Award from the National Center for Advancing Translational Research (KL2 TR 003097).

We would also like to thank the University of Alabama at Birmingham/Lakeshore Foundation Research Collaborative. The development of this project would not have been possible without the collaboration of various experts within the collaborative.

Authors' Contributions

JW, BL, and MC created the initial manuscript draft. BL, MC, JG, and JW developed the teleassessment protocol. H-JY, TM, MT, and JR designed the study procedures related to the teleassessments. All authors contributed equally to later manuscript drafts.

Conflicts of Interest

None declared.

References

1. Fox MH, Witten MH, Lullo C. Reducing Obesity Among People With Disabilities. *J Disabil Policy Stud* 2014 Dec;25(3):175-185 [FREE Full text] [doi: [10.1177/1044207313494236](https://doi.org/10.1177/1044207313494236)] [Medline: [26113785](https://pubmed.ncbi.nlm.nih.gov/26113785/)]
2. Froehlich-Grobe K, Lee J, Washburn RA. Disparities in obesity and related conditions among Americans with disabilities. *Am J Prev Med* 2013 Jul;45(1):83-90. [doi: [10.1016/j.amepre.2013.02.021](https://doi.org/10.1016/j.amepre.2013.02.021)] [Medline: [23790992](https://pubmed.ncbi.nlm.nih.gov/23790992/)]
3. Rimmer JH, Wang E, Yamaki K, Davis B. Documenting Disparities in Obesity and Disability. FOCUS Technical Brief. 2010. URL: https://ktdrr.org/ktlibrary/articles_pubs/ncddrwork/focus/focus24/ [accessed 2021-10-25]
4. Carroll DD, Courtney-Long EA, Stevens AC, Sloan ML, Lullo C, Visser SN, et al. Vital signs: disability and physical activity--United States, 2009-2012. *MMWR Morb Mortal Wkly Rep* 2014 May 09;63(18):407-413 [FREE Full text] [Medline: [24807240](https://pubmed.ncbi.nlm.nih.gov/24807240/)]
5. Rimmer JH, Chen M, Hsieh K. A conceptual model for identifying, preventing, and managing secondary conditions in people with disabilities. *Phys Ther* 2011 Dec;91(12):1728-1739. [doi: [10.2522/ptj.20100410](https://doi.org/10.2522/ptj.20100410)] [Medline: [22003160](https://pubmed.ncbi.nlm.nih.gov/22003160/)]
6. Martin GKA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. *Health Psychol Rev* 2016 Dec;10(4):478-494. [doi: [10.1080/17437199.2016.1198240](https://doi.org/10.1080/17437199.2016.1198240)] [Medline: [27265062](https://pubmed.ncbi.nlm.nih.gov/27265062/)]
7. Lai B, Wilroy J, Young H, Howell J, Rimmer JH, Mehta T, et al. A mobile app to promote adapted exercise and social networking for people with physical disabilities: usability study. *JMIR Form Res* 2019 Mar 19;3(1):e11689 [FREE Full text] [doi: [10.2196/11689](https://doi.org/10.2196/11689)] [Medline: [30888325](https://pubmed.ncbi.nlm.nih.gov/30888325/)]
8. Rimmer JH, Mehta T, Wilroy J, Lai B, Young H, Kim Y, et al. Rationale and design of a Scale-Up Project Evaluating Responsiveness to Home Exercise And Lifestyle Tele-Health (SUPER-HEALTH) in people with physical/mobility disabilities: a type 1 hybrid design effectiveness trial. *BMJ Open* 2019 Mar 30;9(3):e023538 [FREE Full text] [doi: [10.1136/bmjopen-2018-023538](https://doi.org/10.1136/bmjopen-2018-023538)] [Medline: [30928927](https://pubmed.ncbi.nlm.nih.gov/30928927/)]
9. Lai B, Chiu C, Pounds E, Tracy T, Mehta T, Young H, et al. COVID-19 modifications for remote teleassessment and teletraining of a complementary alternative medicine intervention for people with multiple sclerosis: protocol for a randomized controlled trial. *JMIR Res Protoc* 2020 Jul 03;9(7):e18415 [FREE Full text] [doi: [10.2196/18415](https://doi.org/10.2196/18415)] [Medline: [32540838](https://pubmed.ncbi.nlm.nih.gov/32540838/)]
10. Rubin J, Chisnel D. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. Indianapolis, IN: John Wiley & Sons; 2008.
11. Billingham SAM, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. *BMC Med Res Methodol* 2013;13:104 [FREE Full text] [doi: [10.1186/1471-2288-13-104](https://doi.org/10.1186/1471-2288-13-104)] [Medline: [23961782](https://pubmed.ncbi.nlm.nih.gov/23961782/)]
12. Godin Leisure-Time Exercise Questionnaire. *Medicine and Science in Sports and Exercise* 1997;29(Supplement):36-38. [doi: [10.1097/00005768-199706001-00009](https://doi.org/10.1097/00005768-199706001-00009)]
13. Motl RW, Snook EM, McAuley E, Scott JA, Douglass ML. Correlates of physical activity among individuals with multiple sclerosis. *Ann Behav Med* 2006 Oct;32(2):154-161. [doi: [10.1207/s15324796abm3202_13](https://doi.org/10.1207/s15324796abm3202_13)] [Medline: [16972813](https://pubmed.ncbi.nlm.nih.gov/16972813/)]

14. Vasudevan V, Rimmer JH, Kviz F. Development of the barriers to physical activity questionnaire for people with mobility impairments. *Disabil Health J* 2015 Oct;8(4):547-556 [FREE Full text] [doi: [10.1016/j.dhjo.2015.04.007](https://doi.org/10.1016/j.dhjo.2015.04.007)] [Medline: [26087721](https://pubmed.ncbi.nlm.nih.gov/26087721/)]
15. Wójcicki TR, White SM, McAuley E. Assessing outcome expectations in older adults: the multidimensional outcome expectations for exercise scale. *J Gerontol B Psychol Sci Soc Sci* 2009 Jan 29;64(1):33-40 [FREE Full text] [doi: [10.1093/geronb/gbn032](https://doi.org/10.1093/geronb/gbn032)] [Medline: [19181688](https://pubmed.ncbi.nlm.nih.gov/19181688/)]
16. Chiu C, Motl RW, Ditchman N. Validation of the social provisions scale in people with multiple sclerosis. *Rehabil Psychol* 2016 Aug;61(3):297-307 [FREE Full text] [doi: [10.1037/rep0000089](https://doi.org/10.1037/rep0000089)] [Medline: [27088267](https://pubmed.ncbi.nlm.nih.gov/27088267/)]
17. Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al. PROMIS adult health profiles: efficient short-form measures of seven health domains. *Value Health* 2019 May;22(5):537-544. [doi: [10.1016/j.jval.2019.02.004](https://doi.org/10.1016/j.jval.2019.02.004)] [Medline: [31104731](https://pubmed.ncbi.nlm.nih.gov/31104731/)]
18. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. *J Med Internet Res* 2006 Nov;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/)]
19. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol* 2010;1-10 [FREE Full text] [doi: [10.1186/1471-2288-10-1](https://doi.org/10.1186/1471-2288-10-1)] [Medline: [20053272](https://pubmed.ncbi.nlm.nih.gov/20053272/)]
20. Wilroy JD, Martin Ginis KA, Rimmer JH, Wen H, Howell J, Lai B. An E-Learning Program for Increasing Physical Activity Associated Behaviors Among People with Spinal Cord Injury: Usability Study. *JMIR Form Res* 2019 Aug 21;3(3):e14788 [FREE Full text] [doi: [10.2196/14788](https://doi.org/10.2196/14788)] [Medline: [31436161](https://pubmed.ncbi.nlm.nih.gov/31436161/)]
21. Huetteman E. Covid Vaccine Hesitancy Drops Among All Americans, New Survey Shows. *Kaiser Health News*. 2021 Mar 30. URL: <https://khn.org/news/article/covid-vaccine-hesitancy-drops-among-americans-new-kff-survey-shows/> [accessed 2021-10-25]
22. Madhavan S, Sivaramakrishnan A, Bowden MG, Chumbler NR, Field-Fote EC, Kesar TM. Commentary: Remote assessments of gait and balance - Implications for research during and beyond Covid-19. *Top Stroke Rehabil* 2021 Feb 18:1-8. [doi: [10.1080/10749357.2021.1886641](https://doi.org/10.1080/10749357.2021.1886641)] [Medline: [33596774](https://pubmed.ncbi.nlm.nih.gov/33596774/)]
23. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. *N Engl J Med* 2017 Oct 19;377(16):1585-1592. [doi: [10.1056/NEJMs1503323](https://doi.org/10.1056/NEJMs1503323)] [Medline: [29045204](https://pubmed.ncbi.nlm.nih.gov/29045204/)]
24. Lai B, Cederberg K, Vanderbom KA, Bickel CS, Rimmer JH, Motl RW. Characteristics of adults with neurologic disability recruited for exercise trials: a secondary analysis. *Adapt Phys Activ Q* 2018 Oct 01;35(4):476-497. [doi: [10.1123/apaq.2017-0109](https://doi.org/10.1123/apaq.2017-0109)] [Medline: [30382753](https://pubmed.ncbi.nlm.nih.gov/30382753/)]
25. Lai BW, Rimmer JH, Yates A, Jeter A, Young H, Thirumalai M, et al. Critical factors influencing the decision to enroll in a physical activity intervention among a predominant group of adults with spinal cord injury: a grounded theory study. *Spinal Cord* 2021 Jan;59(1):17-25 [FREE Full text] [doi: [10.1038/s41393-020-0530-7](https://doi.org/10.1038/s41393-020-0530-7)] [Medline: [32747672](https://pubmed.ncbi.nlm.nih.gov/32747672/)]

Abbreviations

IRB: Institutional Review Board

M2M: movement-to-music

SUPER-HEALTH: Scale-Up Project Evaluating Responsiveness to Home Exercise And Lifestyle Tele-Health

Edited by G Eysenbach; submitted 21.04.21; peer-reviewed by A Hammons, S Teng, L Sheehy, A Wattanapisit, M Blanco Diaz; comments to author 29.05.21; revised version received 01.07.21; accepted 18.09.21; published 18.11.21.

Please cite as:

Wilroy J, Lai B, Currie M, Young HJ, Thirumalai M, Mehta T, Giannone J, Rimmer J

Teleassessments for Enrollment of Adults With Physical or Mobility Disability in a Home-Based Exercise Trial in Response to COVID-19: Usability Study

JMIR Form Res 2021;5(11):e29799

URL: <https://formative.jmir.org/2021/11/e29799>

doi: [10.2196/29799](https://doi.org/10.2196/29799)

PMID: [34792477](https://pubmed.ncbi.nlm.nih.gov/34792477/)

©Jereme Wilroy, Byron Lai, Madison Currie, Hui-Ju Young, Mohanraj Thirumalai, Tapan Mehta, John Giannone, James Rimmer. Originally published in *JMIR Formative Research* (<https://formative.jmir.org/>), 18.11.2021. 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

Integrating SMS Text Messages Into a Preventive Intervention for Postpartum Depression Delivered via In-Home Visitation Programs: Feasibility and Acceptability Study

Alinne Z Barrera¹, PhD; Jaime Hamil², MPH; Darius Tandon², PhD

¹Department of Psychology, Palo Alto University, Palo Alto, CA, United States

²Institute for Public Health and Medicine, Center for Community Health, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:

Alinne Z Barrera, PhD
Department of Psychology
Palo Alto University
1791 Arastradero Road
Palo Alto, CA, 94304
United States
Phone: 1 6504333854
Email: abarrera@paloaltou.edu

Abstract

Background: The Mothers and Babies (MB) Course is recognized by the US Preventive Services Task Force as an evidence-based preventive intervention for postpartum depression (PPD) that should be recommended to pregnant women at risk for PPD.

Objective: This report examines the feasibility and acceptability of enhancing the MB 1-on-1 intervention by adding 36 SMS text messages that target 3 areas: reinforcement of skills, between-session homework reminders, and responding to self-monitoring texts (ie, MB Plus Text Messaging [MB-TXT]).

Methods: In partnership with 9 home visiting programs, 28 ethnically and racially diverse pregnant women (mean 25.6, SD 9.0 weeks) received MB-TXT. Feasibility was defined by home visitors' adherence to logging into the HealthySMS platform to enter session data and trigger SMS text messages within 7 days of the in-person session. The acceptability of MB-TXT was measured by participants' usefulness and understanding ratings of the SMS text messages and responses to the self-monitoring SMS text messages.

Results: On average, home visitors followed the study protocol and entered session-specific data between 5.50 and 61.17 days following the MB 1-on-1 sessions. A high proportion of participants responded to self-monitoring texts (25/28, 89%) and rated the text message content as very useful and understandable.

Conclusions: This report contributes to a growing body of research focusing on digital adaptations of the MB course. SMS is a low-cost, accessible digital tool that can be integrated into existing interventions. With appropriate resources to support staff, it can be implemented in community-based organizations and health care systems that serve women at risk for PPD.

Trial Registration: ClinicalTrials.gov NCT03420755; <https://clinicaltrials.gov/ct2/show/NCT03420755>

(*JMIR Form Res* 2021;5(11):e30995) doi:[10.2196/30995](https://doi.org/10.2196/30995)

KEYWORDS

perinatal mental health; postpartum depression; public health; SMS; technology

Introduction

Background

Women are at an increased risk of psychological distress during the perinatal period (ie, pregnancy and the first year post partum), with low-income women disproportionately affected.

Regardless of the ethnic and racial background, low-income women report a 20% prevalence of postpartum depression (PPD) [1,2]. PPD is often preceded by symptoms of depression that are not sufficiently elevated to meet the criteria for a diagnosable disorder, but they negatively impact women and raise their risk for perinatal mood and affective disorders. When considering subthreshold depressive symptoms, the rate among low-income

women is as high as 30% to 45% [3], regardless of factors that have been demonstrated to impact depression, such as race and ethnicity [4]. Unfortunately, perinatal depression is often undetected and untreated [5] despite mandates by professional organizations to regularly screen perinatal women for depression [6] and evidence to support the effective prevention and treatment of these disorders [7].

Multiple factors contribute to the disparity in psychological care received by perinatal women who suffer from depression. Providers, mainly obstetricians, often lack the time, knowledge, and training to appropriately identify women who meet the criteria for diagnosable depression or exhibit subthreshold symptoms [8,9]. It is hoped that this deficit in perinatal mental health care will improve given recent mandates expressed by the American College of Obstetricians and Gynecologists and the United States Preventive Services Task Force (USPSTF), with both organizations releasing strong statements recommending that women be regularly screened for depression during the perinatal period [6,10]. In 2019, the USPSTF took one step closer to improving the quality of mental health care for perinatal women when they provided evidence and support for the prevention of PPD interventions [11]. Of major significance was their statement that at-risk perinatal women exhibiting subthreshold symptoms or other risk factors of developing PPD should be referred to receive psychological interventions aimed at prevention of PPD and the subsequent negative consequences [12].

Perinatal women report barriers that interfere with their ability to advocate for their mental health needs, often expressing a lack of knowledge regarding psychological symptoms or help-seeking resources [13], as well as mental illness stigma or not fulfilling the role of motherhood [14,15]. Lack of transportation or childcare is a common barrier among women from low-income and ethnically diverse backgrounds [16,17]. Barriers to mental health care can be addressed through system-wide changes to how maternal mental health services are designed and implemented [17], such as through home visiting (HV) services where trained individuals meet regularly in-home, to assess family needs and provide services or referrals related to maternal and child health, parenting practices, child development, and school readiness. Similarly, there is a growing emphasis on integrating digital aspects into mental health services and interventions with growing evidence for their effectiveness among the general population (eg, Andrews et al, [18]). In fact, standalone and adjunct SMS programs for physical health are effective in promoting behavior change [19] and improving disease management, treatment compliance, and appointment attendance [20].

However, few digital tools have been adapted and tested to address the needs of perinatal women. This is surprising given the ubiquity of mobile devices and growing interest among perinatal women to access maternal health information on their digital devices [21,22]. Recent reports have supported the acceptability of programs and interventions for PPD that are web- or computer-based (eg, Ashford et al and Lau et al [23,24]), use short messaging service (SMS or text messaging) [25,26], and are designed within mobile apps; Hughson et al [27]. Most published digital programs for PPD are in the

protocol, feasibility, or acceptability phases of development [28]. Regardless, in all instances, few low-income and culturally diverse women have taken part in these studies [29].

This study builds on a body of work focused on the implementation and testing of the Mothers and Babies (MB) Course, a PPD preventive intervention that is based on cognitive-behavioral therapy and social learning principles [30]. MB was 1 of only 2 psychological interventions recommended by the USPSTF for prevention of PPD based on a systematic review of 50 empirical studies [12]. MB aims to teach pregnant women skills to cope with changes in their mood and build healthy relationships with supportive adults and their newborn babies. MB has been examined in several controlled prevention trials, including as a face-to-face group intervention with high-risk Latina pregnant women [30,31], as a group intervention to augment HV programs [32-34], and online as a fully automated intervention, Mothers and Babies Online Course (eMB), that included an international sample of English- and Spanish-speaking pregnant women [13]. In-person trials have demonstrated that MB is effective at preventing the onset of PPD and reducing depressive symptoms when compared with usual care. While eMB did not demonstrate a statistical effect in preventing PPD, pregnant women with elevated depressive symptoms reported better depression outcomes if they engaged in eMB relative to those randomized to the information-only condition [35].

Despite the positive outcomes and the unique benefits of the MB modalities, difficulties with engagement and retention occurred across the MB modalities. Time commitment, transportation, childcare, and technology access uniquely impacted the effectiveness of each MB modality (group, 1-on-1, and web-based). For instance, the eMB provided participants with flexibility to engage with the course at their leisure and eliminated the need for transportation and childcare. However, the automated design and advances in technology have resulted in outdated methods to access or engage with web-based interventions. To address this challenge, a recent adaptation of the MB to an SMS platform [36] is in line with its widespread use and reliance on mobile devices. Similarly, the development of MB 1-on-1 intervention delivered individually to a single client as part of a HV program [37] reduces participant burden related to time constraints and childcare needs associated with the group modality. However, the home visitors' ability to ensure immediate and long-term retention of skills taught during the MB 1-on-1 as well as between-session practicing of skills taught remains difficult to ascertain.

Objectives

To maximize the benefits of more personalized and accessible adaptations of the MB intervention, experts in the delivery of MB partnered to create and examine the combination of MB 1-on-1 with SMS enhancements (ie, MB Plus Text Messaging [MB-TXT]). A small quasi-experimental study was conducted in which women were randomized to either MB 1-on-1 or MB-TXT (ie, MB 1-on-1 plus text messaging). The purpose of this report is to determine the feasibility and acceptability of MB-TXT across HV programs. We anticipated that the addition of SMS text messages to the MB-1-on-1 intervention would be

acceptable, as measured by participants' understanding of and engagement with the MB-TXT. The home visitors' adherence to the MB-TXT protocol was also measured and is described in this report as part of examining how feasible the SMS enhancements were. A description of the full trial, including the psychological outcomes, is presented elsewhere.

Methods

Setting

The research team partnered with 9 HV programs in the Midwest to recruit women at high risk of poor pregnancy and parenting outcomes. Participant recruitment was conducted between June 2017 and July 2018. Data from the MB-TXT arm of this study

are presented in this report. The Northwestern University Institutional Review Board approved all study procedures (STU00203918).

Participants

The MB-TXT participants included 28 pregnant women (mean 25.6, SD 9.0 weeks) between the ages of 17 and 31 years (mean 23.7, SD 4.3 years) who mostly self-identified as Black or African American (12/28, 43%) or Hispanic or Latinx (10/28, 36%), and English-speaking (23/28, 82%). Participants were mostly single (16/28, 57%) or living with a partner (8/28, 29%) and stated that their pregnancy was unplanned (22/28, 79%). Demographic information and professional background of the home visitors were not collected as part of this study. See [Table 1](#) for additional participant information.

Table 1. Participant characteristics (N=28).

Demographic characteristics	Value
Age (years), mean (SD)	23.7 (4.3)
Race or ethnicity, n (%)	
Black or African American	12 (43)
Hispanic or Latinx	10 (36)
White	4 (14)
Biracial	1 (4)
Prefer to not specify	1 (4)
Marital status, n (%)	
Single	16 (57)
Living with partner	8 (28)
Engaged	2 (7)
Married	2 (7)
Education, highest level, n (%)	
Grades 1-8	1 (4)
Grades 9-12 (no diploma)	8 (28)
High-school diploma or GED ^a	7 (25)
Some college (no degree)	11 (39)
College degree	1 (4)
Unemployed, n (%)	15 (54)
Annual household income (US \$), n (%)	
0-24,999	22 (81)
25,000-49,999	2 (7)
50,000-74,999	3 (11)
English language, n (%)	
Primarily spoken at home	23 (82)
Pregnancy characteristics	
Weeks pregnant, mean (SD)	25.6 (9.0)
Unplanned pregnancy, n (%)	22 (78)
First childbirth, n (%)	13 (46)

^aGED: General Educational Development test.

Intervention

MB 1-on-1 comprises 12 in-person sessions based on the original group MB [30], each lasting 15 to 20 minutes and delivered by trained home visitors. MB 1-on-1 is divided into 3 modules that correspond to key cognitive-behavioral elements: pleasant activities, thoughts (cognitions), and contact with others (social support). Throughout the MB, mood management skills are integrated with psychoeducational activities that encourage participants to understand the influence of their mood and cognitive-behavioral therapy components. The content is tailored to the specific needs and issues related to the perinatal period.

The MB-TXT includes the delivery of MB 1-on-1 intervention along with an added set of 3 SMS text messages after each of the 12 sessions (see Table 2). SMS text messages were designed to be triggered (ie, sent) following the completion of an MB 1-on-1 session. AZB developed 36 SMS text messages based on an earlier SMS version of the MB [36] to complement the MB 1-on-1 content. The MB 1-on-1 SMS text messages for each intervention session included 1 message focused on each of the following 3 areas: (1) skill reinforcement (eg, *some thoughts just come to us, but we can make a conscious effort to think of positive thoughts*); (2) homework reminders (eg, *what pleasant activity will you try today?*), and (3) self-monitoring (eg, *on a scale of 1-9, how would you rate your mood?*).

Table 2. Mothers and Babies Plus Text Messaging (MB-TXT) examples classified by Mothers and Babies (MB) 1-on-1 session topics.

Topic and HV ^a #	Skill reinforcement	Homework reminder	Self-monitoring
Course introduction			
1	Everybody has stress. It affects how you feel and how you interact with your baby. Do something today to manage stress. It can be as small as counting to 10	PP ^b : Watch and make notes about the 15 min video on being your child's first teacher [38]	How stressed do you feel when you think about becoming or being a mom? Reply 9 to 1 (with 9 being the most stressed and 1 the least stressed)
2	Inner reality: your thoughts. Outer reality: what you do, who you relate to, what is around you. Stop and notice your realities	PP: Remember to keep track of how you are feeling by circling your mood rating on worksheet 2.2 at the end of the day	Reply and let us know what your mood is today on a scale of 1-9 (with 9 being the best day ever, 5 being average, and 1 worst day ever)
Pleasant activities			
3	It's hard to do something pleasant when you are stressed or feeling sad. One pleasant activity can lead to more. Pick one thing you will do to destress	PP: What do you enjoy doing by yourself, with others, and with your baby? Think about it or write it down on worksheet 3.2	Are you working on creating your list of pleasant activities you like to do? Reply Y or N
4	Pleasant activities can be low to no cost, brief, and things that are part of our daily routines	PP: Do not forget to do 1 pleasant activity this week. Fit it into your week by planning a day/time to do it using worksheet 4.3	Were you able to schedule or complete a pleasant activity? Reply Y/N (If Y, also text us what you did!)
5	Babies learn by playing. Doing pleasant activities with you enhances their development. Notice how and what other babies in your community like to play	PP: Remember to keep track of how you are feeling (Quick Mood Scale) and how many pleasant activities you do each day on worksheet 5.3	Reply and let us know your mood for today on a scale of 1 to 9 (with 9 being the best)
Thoughts			
6	People have helpful and harmful thoughts and both types affect how we feel. Make an effort to think helpful thoughts or imagine positive images	PP: What helpful or harmful thoughts have you had this week about being pregnant or becoming a mom? Write them down on worksheet 6.6	Text us a positive thought you have about becoming or being a mom
7	Talking back to harmful thoughts is one way to reduce them. Today, change one of your harmful thoughts into a more helpful thought	PP: Keep track of your mood (Quick Mood Scale). Count how many helpful/harmful thoughts you have this week. Decrease harmful thoughts with a new skill on worksheet 7.4	Have you noticed any harmful thoughts you have? Reply Y/N (In addition, tell us if you used a strategy to reduce it.)
8	Your thoughts can affect your future and your baby's future. It is helpful to have thoughts about your future so you can act in ways to achieve your goals	PP: This week think about what you want your baby's ideal future to look like. What can you begin to do now to prepare for it? Put your ideas on worksheet 8.3	Reply and let us know your mood for today on a scale of 1-9 (with 9 being the best)
Contact with others			
9	Being with others can affect how you feel. Many people feel sad or depressed when they do not have positive contacts or have negative contacts with others	PP: On worksheet 9.4 rate how you are feeling and the number of positive and negative contacts you have each day. Try to do this every day of this week	Tell us about one positive contact you had this week. Text us who you were with and what you did together?
10	Support from others is important for all of us. Positive contacts help you when life gets hard. Different people can provide support for different things	PP: Knowing who you can count on to support your baby is this week's personal project. Write down who those people are on worksheet 10.3	Think about who supported you this week if you were having a difficult day. Text us what kind of support they gave you

Topic and HV ^a #	Skill reinforcement	Homework reminder	Self-monitoring
11	It is important to have your needs met and express what you need. Positive, clear, and direct requests are the best way to communicate. Practice being assertive	PP: This week try to ask for something from anyone—even your home visitor! What way of asking will help get your needs met—passive, aggressive, and assertive?	Reply and let us know your mood for today on a scale of 1 to 9 (with 9 being the best)
Planning for the future			
12	Creating a healthy reality for you and your baby must do with your activities, your thoughts, and your contact with others	Congratulations on finishing the MB! On a scale of 1 to 9, how helpful were texts in reminding you about main ideas and personal projects (9 is most helpful)?	Congratulations on finishing the MB! On a scale of 1 to 9, how helpful were texts in reminding you about main ideas and personal projects (9 is most helpful)?

^aHV: home visiting.

^bPP: personal project.

Measures

Demographic information included age, race or ethnicity, language spoken at home, country of birth and years in the United States, marital status, education, employment, and annual income. Pregnancy history assessed for pregnancy planning, pregnancy length, and previous childbirth.

Feasibility of MB-TXT

MB-TXT feasibility was defined by HV adherence to the MB-TXT protocol, which included postsession documentation of the participant session data (ie, date when the MB 1-on-1 session was conducted), whether the personal project for each MB 1-on-1 session was completed [yes/no], text message ratings for usefulness and understanding, and optional notes), and the triggering of SMS text messages within 7 days of completing each in-person MB 1-on-1 session.

All self-monitoring SMS text messages invited a response from the participants and were used as a proxy measure of engagement and readability of the SMS text messages. Six of the 12 self-monitoring SMS text messages invited participants to provide a numerical rating of their stress (session 1; 9=most stressed to 1=least stressed), their mood (sessions 2, 5, 8, and 11; 9=best day ever to 1=worst day ever), or overall helpfulness of the SMS text messages (session 12; 9=most helpful to 1=least helpful). The remaining 6 self-monitoring SMS text messages invited participants to respond with a message describing how they applied the intervention skills taught during the MB 1-on-1 sessions.

Acceptability of MB-TXT

Participants' perceived utility and comprehension of the SMS text messages was measured using a 4-point Likert scale that assessed the usefulness (ie, 4=very useful to 1=not at all useful) and understanding (ie, 4=totally understood to 1=did not understand at all) of each SMS text message. Participant engagement in the MB-TXT was based on the number of SMS text messages sent by the participant in response to the self-monitoring texts that prompted a response (eg, *Are you working on creating your list of pleasant activities you like to do? Reply Y or N*).

SMS Platform

HealthySMS [39] is a web-based platform designed to send health-related SMS text messages. The platform includes a dashboard that allows administrators to set up users as individual recipients or within groups, to trigger (ie, initiate) texts, to view planned activities (eg, upcoming messages) or generate data (eg, messages received by users), and to track project-specific information, such as home visitor entries of MB 1-on-1 session data. Home visitors received formal training on the MB 1-on-1 intervention before completing the training on the HealthySMS platform.

Design

Participants were referred by their HV programs based on the study eligibility criteria, which included (1) being pregnant, (2) English or Spanish speaking, (3) >16 years, (4) and the ability to receive SMS text messages. Home visitors in each HV program identified and referred clients to the research team; research team members followed up with participants by telephone to confirm eligibility and to obtain informed consent. Before any study activities, informed consent was obtained from all participants, including clients and home visitors. Informed consent was obtained via Research Electronic Data Capture (REDCap) [40,41] for web-based consent, and verbal consent was obtained over the phone (yes or I agree) and documented by a research team member. All participants were mailed a copy of their consent form. A waiver of consent for parents or guardians of participants under 18 years of age was granted by the Northwestern University Institutional Review Board. In the US state where the study was conducted, parent or guardian consent was not required for a client under 18 years to enroll in a HV program. Participants had the option to opt out of the study activities by notifying the study team at any time.

After completing each of the MB 1-on-1 sessions, home visitors were instructed to log-in to the HealthySMS platform to enter session data and trigger session-specific SMS text messages. Once triggered, the asynchronous SMS text messages for each MB 1-on-1 session were automatically sent in a linear and sequential order, at different times of the day within a 12-hour block of time each day, 36-72 hours apart.

Results

Feasibility

Feasibility was defined by home visitors' adherence to the MB-TXT protocol, which instructed them to enter session data into the HealthySMS platform within 7 days of each of the MB 1-on-1 sessions. MB-TXT SMS text messages were triggered, on average, 26.18 (SD 14.40) days following the MB 1-on-1 session (range 5.50-61.17 days).

Response Rates

Overall, 68% (19/28) of participants completed the MB 1-on-1 intervention and received up to 36 MB-TXT SMS text messages. Of the 32% (9/28) remaining participants, 28% (8/28) did not

complete the MB 1-on-1 intervention (ie, received fewer than 36 MB-TXT SMS text messages owing to dropout or premature termination). The self-monitoring SMS text messages were the only SMS text message type used to invite a participant response (Table 3). Almost all participants (25/28, 89%) responded to at least one of the 12 self-monitoring MB-TXT SMS text messages, with all participants in the sample responding to an average of 8.52 (SD 6.35) text messages. Of the 11% (3/28) of participants who did not respond to any of the self-monitoring SMS text messages, 7% (2/28) were foreign-born and 67% (19/28) were Spanish-speakers. Of note, however, these 7% (2/28) of participants did not complete the MB 1-on-1 intervention. No other group differences emerged between those who did and did not respond to the self-monitoring SMS text messages.

Table 3. Participant responses to Mothers and Babies Plus Text Messaging stress and mood rating text messages (N=28).

Session	Text message	Response rate, n (%)	Mean (SD)	Range	Narrative responses
1	How stressed do you feel when you think about becoming or being a mom? Reply 9 to 1 (with 9 being the most stressed and 1 the least stressed)?	16 (57)	3.64 (2.44)	1-8	"Because I have to work out whether or not I'm going to go back too [sic] work after giving birth or sit it out for a few months...and putting the new babysitting plan in action if I decide to go back to work; It feels good. I love being a mom, my kids, my life"
2	Reply and let us know your mood for today on a scale of 1-9 (with 9 being the best day ever, 5 being average, and 1 worst day ever).	18 (66)	7.06 (1.43)	4-9	"Today my son graduated from kindergarten. I'm so proud and excited!"
5	Reply and let us know your mood for today on a scale of 1-9 (with 9 being the best).	17 (61)	7.5 (1.09)	6-9	"I feel good and am present"
8	Reply and let us know your mood for today on a scale of 1-9 (with 9 being the best).	19 (68)	7.59 (1.77)	2-9	"Just got back from _____ [and the] train was 2 hours late—they lost my luggage"
11	Reply and let us know your mood for today on a scale of 1-9 (with 9 being the best).	15 (55)	7.18 (2.14)	2-9	N/A ^a

^aN/A: not applicable.

Consistent with the MB 1-on-1 content, the self-monitoring SMS texts invited participants to respond with a numerical rating of stress (session 1; 9=most stressed to 1=least stressed) and mood (sessions 2, 5, 8, and 11; 9=best day ever to 1=worst day ever). All participants received the session 1 self-monitoring text message, with 57% (16/28) responding to these SMS text messages with an average stress rating of 3.64 (SD 2.44). MB-TXT mood ratings (for MB 1-on-1 sessions 2, 5, 8, and 11) were high, ranging from 7.06 to 7.59 out of 9. Session 12 self-monitoring SMS text message assessed the overall helpfulness of the MB-TXT, with respondents reporting high ratings (mean 8.22, SD 1.30; range 5-9). In 3 instances participants provided a narrative response instead of a numerical rating (eg, *Yes it was* in response to session 12 helpfulness of the MB-TXT) or as context to their numerical rating (eg, "...3, because I have to work out whether or not I'm going to back

too [sic] work after giving birth or sit it out for a few months and putting the new babysitting plan in action if I decide to go back to work" in response to the session 1 inquiry about stress related to motherhood).

Response rates for the self-monitoring SMS text messages that invited a response on how participants were implementing intervention skills ranged from as low as 20% to as high as 65%. In their responses, participants provided examples of the pleasant activities (eg, "Yes, I crocheted on my son's blanket while we listened to music"), positive thoughts (eg, "I must take care of myself so I can take care of those I love"), and type of contacts (eg, "Contact with my newborn. My baby was born") they engaged in (Table 4). Skill reinforcement and homework reminder SMS text message types did not invite a participant response; however, 21 and 18 responses describing participant reactions or reflections to the texts, respectively, were received.

Table 4. Participant responses to Mothers and Babies Plus Text Messaging self-monitoring text messages.

Session	Topic	Text message	Narrative responses
4	<ul style="list-style-type: none"> Pleasant activities What do you like to do? Overcoming obstacles 	Were you able to schedule or complete a pleasant activity? Reply Y/N (If Y, also text us what you did!)	<ul style="list-style-type: none"> “Yes, dancing.” “Yes, I went shopping for Christmas gifts (Spanish)” “Yes, I crocheted on my son’s blanket while we listened to music” “Yes, I was able to take my kids to the park and teach my daughter how to rollerblade and jump rope” “Yes, I went shopping and to a family BBQ”
6	<ul style="list-style-type: none"> Thoughts What are thoughts? Helpful/positive thoughts 	Text us a positive thought you have about becoming or being a mom	<ul style="list-style-type: none"> “I am excited to see my baby’s eyes and smile” “I must take care of myself so I can take care of those I love” “I’m doing a good job! Motherhood is challenging and it [is] ok to struggle!!!” “My son is cute” “I’m looking forward to being an even better mom than before. Watching and growing with my daughter. Being there for both my children. Things are looking up as I become more positive about becoming a parent. I’m even thinking of a career change as a result”
9	<ul style="list-style-type: none"> Contact with others Breaking the cycle between negative mood and fewer positive contacts 	Tell us about one positive contact you had this week. Text us who you were with and what you did together?	<ul style="list-style-type: none"> “On Monday my dad and my brother came to my house we had dinner together and we talk about our childhood” “I had positive contact with my children. We played with blocks, I read stories and we had a great day” “My friend Mara met me at the park and walked back home with me. She walked my dog while I pushed the stroller” “Contact with my newborn. My baby was born”

Acceptability

A total of 293 MB-TXT sets of 3 messages were triggered upon completion of the MB 1-on-1 session. A total of 879 MB-TXT messages were sent through the HealthySMS platform to all 28 participants (mean 31.39, SD 7.98; range 9-36). Overall, 68% (19/28) of participants were sent the full dose of MB-TXT text messages (ie, 684 SMS text messages), given that they attended all 12 MB 1-on-1 sessions. The remaining 32% (9/28) of participants who did not complete all 12 MB 1-on-1 sessions received fewer than 36 MB-TXT SMS text messages, given that they were designed to be triggered postsession. Acceptability assessments indicated that SMS text messages were rated as useful to very useful (range 64%-86.3%) and understandable to totally understandable (range 80%-100%).

Discussion

Principal Findings

The MB was recently recognized by the USPSTF [12] as an evidence-based prevention of PPD intervention that should be recommended for high-risk perinatal women. This report contributes to a growing line of research focused on MB adaptations to meet the needs of perinatal women and the agencies that serve them. This study examined the acceptability and feasibility of the MB 1-on-1 intervention delivered to HV clients with added SMS text messages that reinforced skills taught in each in-person session, reminded participants of between-session homework, and focused on self-monitoring of symptoms and the application of skills to their daily lives. In

addition, self-monitoring SMS text messages, which invited participants to respond, served as a proxy measure of engagement with the MB-TXT program. As expected, participants found the 36 MB-TXT SMS text messages to be highly acceptable and engaged with the program by responding to more SMS text messages than required. In fact, most participants responded to at least one text message, and text messages were rated as both useful and easily understandable, which is critical, especially when aiming to reinforce skills beyond the therapeutic session and attempting to keep the users engaged with the digital approach. Although the high acceptability of the MB-TXT was consistent with previous reports using SMS interventions with perinatal populations, participants in this study engaged at higher rates with the MB-TXT. By comparison, Broom et al [25] reported that 45% of postpartum participants responded to at least one of the treatment adjunct text messages (vs 25/28, 89% of participants in this study) and only 7.30% (vs 55%-68%; 15-19 participants in this study) responded to SMS text messages that allowed responses.

Two key components of the MB prevention intervention are: (1) teaching women how to become aware of their emotional states and related behaviors and (2) practice skills taught in each session with the goal of reducing their risk of PPD. These principles were targeted with MB-TXT self-monitoring and personal project reminder text messages. On the basis of the participants’ engagement with the MB-TXT, about 50% of the women affirmed that they were able to implement the skills taught through MB 1-on-1, a finding that is consistent with

previous reports that used SMS to enhance behavioral interventions [36,42]. Similarly, home visitors indicated that approximately 50% of the women completed session-specific personal projects, which exceeds the rates found in previous MB trials [43].

The success of MB-TXT was largely dependent on the home visitors' engagement with the HealthySMS platform. In this study, the home visitors' level of adherence to accessing the platform within 7 days of the MB 1-on-1 session served as a feasibility measurement. Although most of the home visitors accessed the HealthySMS platform and the participants, as a result, received the full allotted set of MB-TXT SMS text messages, fewer than expected were able to access it within the protocol's timeframe of 7 days. These implementation difficulties were often owing to logistical (eg, time efficiency by *saving* all data entry for a nonfield day), technological (eg, failure to understand how to use the HealthySMS platform), and personal barriers (eg, forgetting to trigger messages within 7 days). Regardless, these noted barriers suggest that home visitors may need additional resources to support their role in MB-TXT implementation. For instance, there were 6 participants for whom research staff needed to regularly assist their home visitors by entering session data into the HealthySMS platform. Four of these participants dropped out of their respective HV programs and ceased to receive the full dose of the MB-TXT. There were also participant and systemic events that likely contributed to challenges in feasibility (eg, disconnected participant telephone numbers or the blocking of the MB-TXT phone number and loss of funding by home visitation agencies and home visitors no longer being affiliated with their agency, respectively).

Limitations

In addition to the noted challenges of implementing the MB-TXT as an enhancement to the MB 1-on-1 intervention,

there are several limitations to consider. The factors that contribute to acceptability ratings remain unknown. That is, the valence of these ratings may be due to external factors not associated with MB-TXT, such as home visitors' understanding of and competency in delivering the MB 1-on-1 or participants' desire to respond favorably to maintain their relationship with the home visitor (ie, social desirability). Related, acceptability ratings were gathered in aggregate by session number and not by text message type. As such, it is unclear whether a specific text message or message type may have been rated differently (eg, for a specific session, the skill reinforcement text may have been experienced as very useful, but the self-monitoring SMS text message may not have been). This lost data point has implications for future iterations of the MB-TXT and the research team's goal of improving the program content and impact on PPD. The wording of SMS text messages may have also influenced how participants engaged in MB-TXT. For instance, the skills reinforcement self-monitoring messages that invited participants to share how they applied intervention skills to their daily lives may have inadvertently discouraged those who did not implement the skills from submitting a response. As such, it is difficult to ascertain whether those who did not respond were participants who, in fact, did implement the skills but failed to respond versus those who may have struggled to implement the skills and did not want to provide that feedback or who simply ignored the invitation to provide a response.

Conclusions

The findings of this study are promising as low-cost evidence-based digital interventions are being developed with the goal of reaching underserved communities of perinatal women at risk for PPD. Future iterations of MB-TXT will aim to integrate lessons learned from this study, especially as they relate to digital interventions that rely on or integrate human support as part of implementation.

Acknowledgments

The authors would like to thank the Home Visiting (HV) programs that collaborated with this project. They would like to thank the research assistants and the students who collaborated with this project. They also thank Chris Karr, who provided the technical support for this project. The authors thank the following funder for the support of this project: National Institute on Minority and Health Disparities: R21 MD011320-01 (SDT).

Conflicts of Interest

None declared.

References

1. Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. *Evid Rep Technol Assess (Summ)* 2005 Feb(119):1-8. [doi: [10.1037/e439372005-001](https://doi.org/10.1037/e439372005-001)] [Medline: [15760246](https://pubmed.ncbi.nlm.nih.gov/15760246/)]
2. Moses-Kolko EL, Roth EK. Antepartum and postpartum depression: healthy mom, healthy baby. *J Am Med Womens Assoc* (1972) 2004;59(3):181-191. [Medline: [15354371](https://pubmed.ncbi.nlm.nih.gov/15354371/)]
3. Howell EA, Mora PA, Horowitz CR, Leventhal H. Racial and ethnic differences in factors associated with early postpartum depressive symptoms. *Obstet Gynecol* 2005 Jun;105(6):1442-1450 [FREE Full text] [doi: [10.1097/01.AOG.0000164050.34126.37](https://doi.org/10.1097/01.AOG.0000164050.34126.37)] [Medline: [15932842](https://pubmed.ncbi.nlm.nih.gov/15932842/)]
4. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. *Obstet Gynecol* 2004 Apr;103(4):698-709. [doi: [10.1097/01.AOG.0000116689.75396.5f](https://doi.org/10.1097/01.AOG.0000116689.75396.5f)] [Medline: [15051562](https://pubmed.ncbi.nlm.nih.gov/15051562/)]

5. Bales M, Pambrun E, Melchior M, Glangeaud-Freudenthal NM, Charles M, Verdoux H, et al. Prenatal psychological distress and access to mental health care in the ELFE cohort. *Eur Psychiatry* 2015 Feb;30(2):322-328. [doi: [10.1016/j.eurpsy.2014.11.004](https://doi.org/10.1016/j.eurpsy.2014.11.004)] [Medline: [25542278](https://pubmed.ncbi.nlm.nih.gov/25542278/)]
6. Siu AL, US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, et al. Screening for depression in adults: US preventive services task force recommendation statement. *JAMA* 2016 Jan 26;315(4):380-387. [doi: [10.1001/jama.2015.18392](https://doi.org/10.1001/jama.2015.18392)] [Medline: [26813211](https://pubmed.ncbi.nlm.nih.gov/26813211/)]
7. Cuijpers P, Brännmark JG, van Straten A. Psychological treatment of postpartum depression: a meta-analysis. *J Clin Psychol* 2008 Jan;64(1):103-118. [doi: [10.1002/jclp.20432](https://doi.org/10.1002/jclp.20432)] [Medline: [18161036](https://pubmed.ncbi.nlm.nih.gov/18161036/)]
8. Byatt N, Biebel K, Debordes-Jackson G, Lundquist RS, Moore Simas TA, Weinreb L, et al. Community mental health provider reluctance to provide pharmacotherapy may be a barrier to addressing perinatal depression: a preliminary study. *Psychiatr Q* 2013 Jun;84(2):169-174. [doi: [10.1007/s11126-012-9236-0](https://doi.org/10.1007/s11126-012-9236-0)] [Medline: [22941573](https://pubmed.ncbi.nlm.nih.gov/22941573/)]
9. Byatt N, Moore Simas TA, Biebel K, Sankaran P, Pbert L, Weinreb L, et al. PRogram In Support of Moms (PRISM): a pilot group randomized controlled trial of two approaches to improving depression among perinatal women. *J Psychosom Obstet Gynaecol* 2018 Dec;39(4):297-306 [FREE Full text] [doi: [10.1080/0167482X.2017.1383380](https://doi.org/10.1080/0167482X.2017.1383380)] [Medline: [28994626](https://pubmed.ncbi.nlm.nih.gov/28994626/)]
10. The American College of Obstetricians and Gynecologists Committee Opinion no. 630. Screening for perinatal depression. *Obstet Gynecol* 2015 May;125(5):1268-1271. [doi: [10.1097/01.AOG.0000465192.34779.dc](https://doi.org/10.1097/01.AOG.0000465192.34779.dc)] [Medline: [25932866](https://pubmed.ncbi.nlm.nih.gov/25932866/)]
11. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Interventions to prevent perinatal depression: US Preventive Services Task Force Recommendation Statement. *JAMA* 2019 Feb 12;321(6):580-587. [doi: [10.1001/jama.2019.0007](https://doi.org/10.1001/jama.2019.0007)] [Medline: [30747971](https://pubmed.ncbi.nlm.nih.gov/30747971/)]
12. O'Connor E, Senger CA, Henninger ML, Coppola E, Gaynes BN. Interventions to prevent perinatal depression: evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2019 Feb 12;321(6):588-601. [doi: [10.1001/jama.2018.20865](https://doi.org/10.1001/jama.2018.20865)] [Medline: [30747970](https://pubmed.ncbi.nlm.nih.gov/30747970/)]
13. Barrera AZ, Wickham RE, Muñoz RF. Online prevention of postpartum depression for Spanish- and English-speaking pregnant women: a pilot randomized controlled trial. *Internet Interv* 2015 Sep 01;2(3):257-265 [FREE Full text] [doi: [10.1016/j.invent.2015.06.002](https://doi.org/10.1016/j.invent.2015.06.002)] [Medline: [26273567](https://pubmed.ncbi.nlm.nih.gov/26273567/)]
14. Fonseca A, Gorayeb R, Canavarró MC. Women's help-seeking behaviours for depressive symptoms during the perinatal period: socio-demographic and clinical correlates and perceived barriers to seeking professional help. *Midwifery* 2015 Dec;31(12):1177-1185. [doi: [10.1016/j.midw.2015.09.002](https://doi.org/10.1016/j.midw.2015.09.002)] [Medline: [26433622](https://pubmed.ncbi.nlm.nih.gov/26433622/)]
15. Thomas LJ, Scharp KM, Paxman CG. Stories of postpartum depression: exploring health constructs and help-seeking in mothers' talk. *Women Health* 2014;54(4):373-387. [doi: [10.1080/03630242.2014.896442](https://doi.org/10.1080/03630242.2014.896442)] [Medline: [24617862](https://pubmed.ncbi.nlm.nih.gov/24617862/)]
16. Alvidrez J, Azocar F. Distressed women's clinic patients: preferences for mental health treatments and perceived obstacles. *Gen Hosp Psychiatry* 1999;21(5):340-347. [doi: [10.1016/s0163-8343\(99\)00038-9](https://doi.org/10.1016/s0163-8343(99)00038-9)] [Medline: [10572775](https://pubmed.ncbi.nlm.nih.gov/10572775/)]
17. Goodman JH. Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. *Birth* 2009 Mar;36(1):60-69. [doi: [10.1111/j.1523-536X.2008.00296.x](https://doi.org/10.1111/j.1523-536X.2008.00296.x)] [Medline: [19278385](https://pubmed.ncbi.nlm.nih.gov/19278385/)]
18. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. *J Anxiety Disord* 2018 Apr;55:70-78 [FREE Full text] [doi: [10.1016/j.janxdis.2018.01.001](https://doi.org/10.1016/j.janxdis.2018.01.001)] [Medline: [29422409](https://pubmed.ncbi.nlm.nih.gov/29422409/)]
19. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. *Am J Prev Med* 2009 Feb;36(2):165-173. [doi: [10.1016/j.amepre.2008.09.040](https://doi.org/10.1016/j.amepre.2008.09.040)] [Medline: [19135907](https://pubmed.ncbi.nlm.nih.gov/19135907/)]
20. 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/)]
21. Chilukuri N, West M, Henderson JL, Lawson S, Ehsanipoor R, Costigan K, et al. Information and communication technology use among low-income pregnant and postpartum women by race and ethnicity: a cross-sectional study. *J Med Internet Res* 2015 Jul 03;17(7):e163 [FREE Full text] [doi: [10.2196/jmir.3916](https://doi.org/10.2196/jmir.3916)] [Medline: [26142162](https://pubmed.ncbi.nlm.nih.gov/26142162/)]
22. Peragallo Urrutia R, Berger AA, Ivins AA, Beckham AJ, Thorp JM, Nicholson WK. Internet use and access among pregnant women via computer and mobile phone: implications for delivery of perinatal care. *JMIR Mhealth Uhealth* 2015 Mar 30;3(1):e25 [FREE Full text] [doi: [10.2196/mhealth.3347](https://doi.org/10.2196/mhealth.3347)] [Medline: [25835744](https://pubmed.ncbi.nlm.nih.gov/25835744/)]
23. Ashford MT, Olander EK, Ayers S. Computer- or web-based interventions for perinatal mental health: a systematic review. *J Affect Disord* 2016 Jun;197:134-146. [doi: [10.1016/j.jad.2016.02.057](https://doi.org/10.1016/j.jad.2016.02.057)] [Medline: [26991368](https://pubmed.ncbi.nlm.nih.gov/26991368/)]
24. Lau A, Barnett M, Stadnick N, Saifan D, Regan J, Wiltsey Stirman S, et al. Therapist report of adaptations to delivery of evidence-based practices within a system-driven reform of publicly funded children's mental health services. *J Consult Clin Psychol* 2017 Jul;85(7):664-675 [FREE Full text] [doi: [10.1037/ccp0000215](https://doi.org/10.1037/ccp0000215)] [Medline: [28471210](https://pubmed.ncbi.nlm.nih.gov/28471210/)]
25. Broom MA, Ladley AS, Rhyne EA, Halloran DR. Feasibility and perception of using text messages as an adjunct therapy for low-income, minority mothers with postpartum depression. *JMIR Ment Health* 2015 Mar 16;2(1):e4 [FREE Full text] [doi: [10.2196/mental.4074](https://doi.org/10.2196/mental.4074)] [Medline: [26543910](https://pubmed.ncbi.nlm.nih.gov/26543910/)]
26. Rhyne EP, Borawski A. Text messaging as an adjunct treatment for urban mothers with postpartum depression. *J Pediatr Health Care* 2014 Nov 1;28(6):e49. [doi: [10.1016/j.pedhc.2014.08.003](https://doi.org/10.1016/j.pedhc.2014.08.003)]

27. Hughson JP, Daly JO, Woodward-Kron R, Hajek J, Story D. The rise of pregnancy apps and the implications for culturally and linguistically diverse women: narrative review. *JMIR Mhealth Uhealth* 2018 Nov 16;6(11):e189 [FREE Full text] [doi: [10.2196/mhealth.9119](https://doi.org/10.2196/mhealth.9119)] [Medline: [30446483](https://pubmed.ncbi.nlm.nih.gov/30446483/)]
28. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. *J Med Internet Res* 2020 Apr 13;22(4):e17011 [FREE Full text] [doi: [10.2196/17011](https://doi.org/10.2196/17011)] [Medline: [32281939](https://pubmed.ncbi.nlm.nih.gov/32281939/)]
29. Lara-Cinisomo S, Ramirez Olarte A, Rosales M, Barrera AZ. A systematic review of technology-based prevention and treatment interventions for perinatal depression and anxiety in Latina and African American women. *Matern Child Health J* 2021 Feb;25(2):268-281. [doi: [10.1007/s10995-020-03028-9](https://doi.org/10.1007/s10995-020-03028-9)] [Medline: [33389589](https://pubmed.ncbi.nlm.nih.gov/33389589/)]
30. Muñoz RF, Le H, Ippen CG, Diaz MA, Urizar GG, Soto J, et al. Prevention of postpartum depression in low-income women: development of the Mamás y Bebés/mothers and babies course. *Cognit Behav Pract* 2007 Feb;14(1):70-83. [doi: [10.1016/j.cbpra.2006.04.021](https://doi.org/10.1016/j.cbpra.2006.04.021)]
31. Le H, Perry DF, Stuart EA. Randomized controlled trial of a preventive intervention for perinatal depression in high-risk Latinas. *J Consult Clin Psychol* 2011 Apr;79(2):135-141. [doi: [10.1037/a0022492](https://doi.org/10.1037/a0022492)] [Medline: [21319897](https://pubmed.ncbi.nlm.nih.gov/21319897/)]
32. Tandon SD, Leis JA, Mendelson T, Perry DF, Kemp K. Six-month outcomes from a randomized controlled trial to prevent perinatal depression in low-income home visiting clients. *Matern Child Health J* 2014 May;18(4):873-881 [FREE Full text] [doi: [10.1007/s10995-013-1313-y](https://doi.org/10.1007/s10995-013-1313-y)] [Medline: [23793487](https://pubmed.ncbi.nlm.nih.gov/23793487/)]
33. Tandon SD, Perry DF, Mendelson T, Kemp K, Leis JA. Preventing perinatal depression in low-income home visiting clients: a randomized controlled trial. *J Consult Clin Psychol* 2011 Oct;79(5):707-712. [doi: [10.1037/a0024895](https://doi.org/10.1037/a0024895)] [Medline: [21806298](https://pubmed.ncbi.nlm.nih.gov/21806298/)]
34. McFarlane E, Burrell L, Duggan A, Tandon D. Outcomes of a randomized trial of a cognitive behavioral enhancement to address maternal distress in home visited mothers. *Matern Child Health J* 2017 Mar;21(3):475-484 [FREE Full text] [doi: [10.1007/s10995-016-2125-7](https://doi.org/10.1007/s10995-016-2125-7)] [Medline: [27535131](https://pubmed.ncbi.nlm.nih.gov/27535131/)]
35. Barrera AZ, Nichols AD. Depression help-seeking attitudes and behaviors among an internet-based sample of Spanish-speaking perinatal women. *Rev Panam Salud Publica* 2015 Mar;37(3):148-153. [Medline: [25988251](https://pubmed.ncbi.nlm.nih.gov/25988251/)]
36. Barrera AZ, Aguilera A, Inlow N, Servin J. A preliminary study on the acceptability of a brief SMS program for perinatal women. *Health Informatics J* 2020 Jun;26(2):1079-1087 [FREE Full text] [doi: [10.1177/1460458219866560](https://doi.org/10.1177/1460458219866560)] [Medline: [31566457](https://pubmed.ncbi.nlm.nih.gov/31566457/)]
37. Darius Tandon S, Leis JA, Ward EA, Snyder H, Mendelson T, Perry DF, et al. Adaptation of an evidence-based postpartum depression intervention: feasibility and acceptability of mothers and babies 1-on-1. *BMC Pregnancy Childbirth* 2018 Apr 11;18(1):93 [FREE Full text] [doi: [10.1186/s12884-018-1726-0](https://doi.org/10.1186/s12884-018-1726-0)] [Medline: [29642868](https://pubmed.ncbi.nlm.nih.gov/29642868/)]
38. My Parents, My Teachers. US Department of Health & Human Services. URL: <https://eclkc.ohs.acf.hhs.gov/video/my-parents-my-teachers> [accessed 2021-11-03]
39. Aguilera A, Bruehlman-Senecal E, Demasi O, Avila P. Automated text messaging as an adjunct to cognitive behavioral therapy for depression: a clinical trial. *J Med Internet Res* 2017 May 08;19(5):e148 [FREE Full text] [doi: [10.2196/jmir.6914](https://doi.org/10.2196/jmir.6914)] [Medline: [28483742](https://pubmed.ncbi.nlm.nih.gov/28483742/)]
40. 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/)]
41. 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/)]
42. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. *Behav Modif* 2015 Nov;39(6):785-804 [FREE Full text] [doi: [10.1177/0145445515595198](https://doi.org/10.1177/0145445515595198)] [Medline: [26187164](https://pubmed.ncbi.nlm.nih.gov/26187164/)]
43. Tandon SD, Ward EA, Hamil JL, Jimenez C, Carter M. Perinatal depression prevention through home visitation: a cluster randomized trial of mothers and babies 1-on-1. *J Behav Med* 2018 Oct;41(5):641-652. [doi: [10.1007/s10865-018-9934-7](https://doi.org/10.1007/s10865-018-9934-7)] [Medline: [29766470](https://pubmed.ncbi.nlm.nih.gov/29766470/)]

Abbreviations

eMB: Mothers and Babies Web-based Course

HV: home visiting

MB: Mothers and Babies

MB-TXT: Mothers and Babies Plus Text Messaging

PPD: postpartum depression

REDCap: Research Electronic Data Capture

USPSTF: United States Preventive Services Task Force

Edited by G Eysenbach; submitted 24.06.21; peer-reviewed by J Osma López, K Heinrich; comments to author 01.09.21; revised version received 28.09.21; accepted 03.10.21; published 18.11.21.

Please cite as:

Barrera AZ, Hamil J, Tandon D

Integrating SMS Text Messages Into a Preventive Intervention for Postpartum Depression Delivered via In-Home Visitation Programs: Feasibility and Acceptability Study

JMIR Form Res 2021;5(11):e30995

URL: <https://formative.jmir.org/2021/11/e30995>

doi: [10.2196/30995](https://doi.org/10.2196/30995)

PMID: [34792478](https://pubmed.ncbi.nlm.nih.gov/34792478/)

©Alinne Z Barrera, Jaime Hamil, Darius Tandon. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 18.11.2021. 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

Informing Content and Feature Design of a Parent-Focused Human Papillomavirus Vaccination Digital Behavior Change Intervention: Synchronous Text-Based Focus Group Study

Elisabeth RB Becker¹, PhD; Ross Shegog¹, PhD; Lara S Savas¹, PhD; Erica L Frost¹, MPH; C Mary Healy², MD; Stanley W Spinner³, MD; Sally W Vernon¹, PhD

¹The University of Texas Health Science Center at Houston, Houston, TX, United States

²Baylor College of Medicine, Houston, TX, United States

³Texas Children's Pediatrics, Houston, TX, United States

Corresponding Author:

Elisabeth RB Becker, PhD

The University of Texas Health Science Center at Houston

7000 Fannin St

Houston, TX, 77030

United States

Phone: 1 713 500 4472

Email: elisabeth.becker@uth.tmc.edu

Abstract

Background: Human papillomavirus (HPV) is a common and preventable sexually transmitted infection; however, vaccination rates in the United States among the target age group, which is 11-12 years, are lower than national goals. Interventions that address the barriers to and facilitators of vaccination are important for improving HPV vaccination rates. Web-based, text-based focus groups are becoming a promising method that may be well suited for conducting formative research to inform the design of digital behavior change intervention (DBCI) content and features that address HPV vaccination decision-making.

Objective: This study aims to explore parental HPV vaccination decision-making processes using a web-based, text-based focus group protocol to inform content and feature recommendations for an HPV prevention DBCI.

Methods: We conducted 4 web-based, text-based synchronous focus groups via Skype with the parents of patients aged 11-13 years within a large urban US pediatric clinic network.

Results: The 22 parents were mostly female, White, non-Hispanic college graduates, and they mostly had private health insurance for their children. Approximately half (14/25, 56%) of the parents' 11-13 year old children had initiated HPV vaccination. Most parents had experience using Skype (19/22, 86%). Approximately half (8/17, 47%) of parents expressed no preference for the focus group format, whereas 47% (8/17) requested a text-only chat format and 6% (1/17) requested an audiovisual format. The three main themes from the qualitative data were barriers to HPV vaccination, facilitators of HPV vaccination, and suggestions for improving the HPV vaccination clinic experience. A total of 11 intervention content and feature recommendations emerged from the themes, including addressing HPV knowledge barriers using trusted sources, designing for a family audience, focusing on the framing of messages, reporting reputable HPV research in a comprehensible format, and expanding the clinic visit experience.

Conclusions: Synchronous text-based focus groups are feasible for conducting formative research on HPV vaccination decision-making. Among well-educated and well-resourced parents, there are barriers such as misinformation and facilitators such as pediatrician recommendations that influence HPV vaccination decision-making. Parents want to conduct their own HPV research as well as receive relevant HPV vaccination advice from their child's pediatrician. In addition, parents want an enhanced clinic visit experience that lets them access and connect to tailored information before and after clinic visits. The results gathered provide guidance for content and features that may inform a more responsive DBCI to address HPV vaccination decision-making among parents.

(JMIR Form Res 2021;5(11):e28846) doi:[10.2196/28846](https://doi.org/10.2196/28846)

KEYWORDS

human papillomavirus; vaccination; qualitative; focus groups; sexually transmitted infection

Introduction

Background

Human papillomavirus (HPV) is the most common sexually transmitted infection (STI) in the United States [1] and worldwide [2]. The US Centers for Disease Control and Prevention (CDC) estimates that approximately 90% of men and 80% of women will be infected with at least one type of HPV in their lives [3]. HPV infection generally occurs within a few months to years of becoming sexually active [4,5].

HPV can cause asymptomatic infections, warts, and cancer in women and men [6]. An estimated 34,800 HPV-attributed cancers of the cervix, vagina, vulva, penis, anus, and oropharynx are diagnosed in the United States every year [7]. Most of these cancers are associated with the infection of HPV types 16 and 18. Although HPV is commonly associated with cervical cancer, its prevalence in oropharyngeal tumors has increased substantially from the 1980s (16%) to the 2000s (73%) [8]. In addition to cancer, HPV causes 300,000 new cases of genital warts each year via HPV types 6 and 11 in the United States [9]. Owing to HPV's prevalence and ease of transmission, there have been global efforts to prevent the spread of HPV infection.

Historically, 3 prophylactic HPV vaccines (2vHPV, 4vHPV, and 9vHPV) have been licensed for use in the United States since the first HPV vaccine became available in 2006 [10,11]. Over time, the dosing regimen, range of protection, and intended patient profile for the HPV vaccine have changed. Since 2016, the 9-valent HPV vaccine (9vHPV) has been the only HPV vaccine sold in the United States [12] and is currently available for males and females aged 9-45 years. The Advisory Committee on Immunization Practices recommends that routine HPV vaccination be initiated for children aged 11-12 years [13]. The Healthy People 2030 goal for HPV vaccination series completion is 80% of adolescents aged 13-15 years. However, national samples of adolescents aged 13-17 years estimate that only 60% have initiated and 40% have completed the HPV vaccination series [14].

The modification of factors that negatively affect parental vaccination decision-making is a key strategy to improving HPV vaccination rates [15,16]. A 2019 systematic review of 41 US-based studies exploring HPV vaccine beliefs found four negative beliefs—perceived adverse effects (ie, promotes sexual activity, too new, and causes illness), perceived lack of necessity (not sexually active), morality concerns (stigmatizing recipients), and skepticism about effectiveness—and one positive belief, that it prevents STIs, across the literature [17]. In comparison, the National Immunization Survey-Teen 2010-2016 trend data found safety concerns, a lack of vaccine knowledge, and perceived lack of necessity as consistent reasons why parents did not initiate HPV vaccination for their adolescents [18]. Among caregivers, the internet is a popular source for information regarding HPV vaccination [19]; however, most websites that contain information about HPV immunization have poor readability [20]. Furthermore, web-based consumer

health information is susceptible to miscommunication, misrepresentation, and misappropriation, which can result in negative health consequences for information seekers and those they care for [21]. HPV prevention digital behavior change interventions (DBCIs) that focus on relatable, understandable, and actionable information may facilitate HPV vaccination decision-making among parents.

Formative research constitutes an important component informing DCBI content and feature design [22-25] and includes understanding the barriers and facilitators of vaccination in the context of HPV. As the adaptation of research protocols to remote formats becomes imperative [26,27], web-based, text-based focus groups may serve as a feasible way to gather DCBI formative data. Web-based synchronous text-based focus groups have addressed a range of issues, including parental attitudes toward the 9-valent HPV vaccine [28], HPV mobile health preferences among young men who have sex with men [29], and views on sex among childhood survivors of cancer [30,31]. Studies comparing web-based synchronous text-based focus groups with the traditional in-person approach have found similar thematic content and quality across the formats [32-34].

Objective

The purpose of this study is to explore parental HPV vaccination decision-making processes within a large urban pediatric clinic network using a web-based, text-based focus group protocol to inform content and feature recommendations for an HPV prevention DCBI. Specifically, we seek to answer the following three research questions:

1. Themes: What are the influences of HPV vaccination decision-making among parents belonging to a large pediatric clinic network?
2. Intervention recommendations: On the basis of the themes, what are the content and feature recommendations for a DCBI to encourage HPV vaccination?
3. Format feasibility: How feasible is conducting formative HPV research with parents using synchronous text-based focus groups?

Methods

Overview

Four 60-minute web-based synchronous text-based focus groups were conducted via Skype (Microsoft Inc) in October 2016 with parents of adolescents aged 11-13 years. The adolescents were patients at a large urban and geographically diverse pediatric clinic network in Texas, United States. Parents were invited to participate in the focus groups through their pediatricians, study recruitment flyers posted in the clinic waiting rooms, and advertisements on the clinic's Facebook page. Recruited parents completed a phone-based screening to assess their eligibility. After verbal consent was received, potential parents were sent a demographic and focus group preference survey. Parents were provided a Skype username and password for the duration of the study along with a Skype guide that included download and

log-in instructions with screenshots. Usernames were created that included the parent's first name and an ID number and were deleted at the conclusion of each session. Parents were instructed to use Skype on a laptop or desktop, so their mobile phone would be available if tech support was needed. Approximately 48 hours before the start of their session, parents were asked to log into their study account and answer the question, "What activities do you like to do with your child or children on the weekend?" as a check that they had successfully downloaded Skype (if needed) and could navigate the chat function. A total of 4 research staff members conducted the synchronous focus groups. This included the lead moderator who posted questions in the group chat, 2 submoderators who took notes and monitored the discussion for opportunities to probe participants, and a technology facilitator who assisted parents with using Skype if needed. A US \$40 web-based gift card incentive was offered to the parents at the completion of their focus group. Study protocols and procedures were approved by the University of Texas Health Science Center at Houston institutional review board (HSC-SPH-15-0202).

Participant Inclusion Criteria

Participants were eligible for the study if they were a parent or legal guardian of a patient aged 11-13 years in the clinic network. Participants needed an internet connection, access to a desktop or laptop with a keyboard, and the ability to download and use the free web-based video conferencing and chat platform, Skype [35]. Participants also needed to be able to read and write in English.

Measures

Demographics

Parents' demographic variables included sex, age, race, ethnicity, education, number of children, child's health insurance status, child's HPV vaccination initiation status, and adolescent vaccine hesitancy status. Parent-adolescent vaccine hesitancy

was assessed using an adapted question from the Parent Attitudes about Childhood Vaccines survey [36], "Overall, how hesitant about adolescent vaccinations (HPV, Tdap [tetanus, diphtheria, and pertussis], Meningitis, and influenza), would you consider yourself to be?" with Likert-scale response options of *not hesitant*, *somewhat hesitant*, *unsure*, *hesitant*, and *very hesitant*.

Skype Use

Skype use was assessed by asking parents how often they use Skype (aware but never use, use sometimes, or use regularly). The preferred Skype focus group format was assessed by asking parents how they would like to communicate for the focus group (Skype chat [text-based only], Skype call [audio and visual], or no preference). Skype logistics were evaluated from the no-show rate for each focus group session and the difficulty in attending the session (ability to log into the account 48 hours before the session, being late to the session, calls with the technology facilitator, and technical difficulties during the session). The costs of using the web-based Skype format were compared against the projected costs of conducting in-person focus groups at the clinics. Although both groups included participation incentive costs, in-person groups would have required additional costs, including transcription, parking reimbursements, and after-hours pay of clinic staff.

HPV Vaccine Decision-Making

Research staff trained in qualitative research conducted the focus groups using a discussion guide comprising questions on parental vaccination attitudes and vaccination decision-making processes (Textbox 1). Topic 3 included a series of DBCI subquestions. The discussion guide was informed by prior qualitative research conducted with the network clinic staff and HPV vaccination barriers and facilitators reported in prior research [28,37-43]. Optional probes were generated beforehand and modified in real time as needed.

Textbox 1. Focus group guide.**Topic 1**

- Prompt: Patients aged 11-12 years typically receive four vaccinations: Tdap, meningococcal, human papillomavirus (HPV), and flu. In the survey you completed, some of you indicated that your adolescent has received all of these vaccinations, whereas others indicated that your adolescent has not received one or more of these vaccinations.
- Question: How do you decide if your adolescent should receive certain vaccinations? Can you please describe that process?

Topic 2

- Prompt: At the network clinics, rates of HPV vaccination are lower than rates of the other adolescent vaccinations. In the survey you completed, some of you said that your adolescent has received the HPV vaccine and some of you said that your adolescent has not yet received the HPV vaccine.
- Question: What were the most important factors in making the decision whether to vaccinate or not vaccinate your adolescent for HPV?

Topic 3

- Prompt: In light of what we have just discussed, let us consider ways to improve the parent experience at the clinic.
- Question: How can the clinic help you as a parent in making decisions around vaccinating your adolescent against HPV?
- Digital behavior change intervention subquestions:
 - How can the clinic best communicate with you about HPV? What channels?
 - Would you like to receive information from the clinic through a phone app? Why or why not?
 - What educational information about the HPV vaccine on a website or phone app would be most helpful to you?

Data Analysis

Demographics were analyzed descriptively, including mean and range for continuous data and frequency and percentage for nominal data. Skype logistics data were analyzed from focus group transcripts and call logs. Cost savings were calculated by comparing the estimated in-person costs of focus groups with the costs of Skype-mediated focus groups. The qualitative analysis was completed in two phases. The first phase involved 4 project staff members creating a preliminary codebook and assessing the frequency of responses to specific questions across the focus groups. The second phase involved the lead moderator (first author, ERBB) conducting a conventional content analysis, the systematic classification process of coding and identifying themes, with coding categories derived directly from the data [44]. In the review of each transcript and accompanying notes, existing codes and broader categories were used to identify themes for parental vaccination decision-making. These findings further informed the content and feature design recommendations of an HPV prevention DBCI [45]. The identified themes were reviewed and retained on the basis of the consensus of the research team.

Results**Demographic Characteristics**

A total of 22 parents with an adolescent aged 11-13 years participated in the web-based, text-based focus groups (Table 1). Parents came from 31% (16/51) of clinics within the network. There were 4-7 parents in each session. Parents were aged 41.9 years (SD 6.1 years); had an average of 2 children; were mostly female (21/22, 95%), White, and non-Hispanic (13/22, 59%); had a graduate or professional degree (10/22, 45%) and had private health insurance for their child or children (18/22, 82%). Most parents were not hesitant (11/22, 50%) or somewhat hesitant (8/22, 36%) toward adolescent vaccinations. Approximately 56% (14/25) of the parents' 11-13 year old children had initiated HPV vaccination. This approximates the demographic characteristics of the clinic network population, where, among children aged 10-17 years, 45% (56,934/127,975) were White and non-Hispanic, 80% (102,223/127,975) had private health insurance, and 58% (74,204/127,975) had initiated the HPV vaccination.

Table 1. Parent demographics (N=22).

Characteristic	Study participants
Parent age (years)	
Value, mean (SD)	41.95 (6.12)
Value, range	30-52
Number of children^a	
Value, mean (SD)	1.95 (1.31)
Value, median (range)	2 (1-5)
Value, mode	1
Total male	19
Total female	24
Parent sex, n (%)	
Male	1 (5)
Female	21 (95)
Parent race and ethnicity, n (%)	
White and non-Hispanic	13 (59)
Black or African American and non-Hispanic	7 (32)
Hispanic ^b	2 (9)
Parent education, n (%)	
High school graduate	4 (18)
College graduate	8 (36)
Graduate or professional degree	10 (45)
Child or children's health insurance status, n (%)	
Private health insurance	18 (82)
Medicaid	2 (9)
State Children's Insurance Program	1 (5)
Military health care	1 (5)
Adolescent vaccination hesitancy status^c, n (%)	
Very hesitant	2 (9)
Hesitant	1 (5)
Unsure	0 (0)
Somewhat hesitant	8 (36)
Not hesitant	11 (50)
Child or children's HPV^d vaccination initiation status^e, n (%)	
9-10 years	
Yes	0 (0)
No	5 (100)
11-13 years	
Yes	14 (56)
No	11 (44)
≥14 years	
Yes	10 (77)
No	3 (23)

Characteristic	Study participants
Skype use, n (%)	
Aware of Skype but never used it	3 (14)
Use sometimes	18 (82)
Use regularly	1 (5)
Preferred focus group format^f, n (%)	
Skype chat (text-based only)	8 (47)
Skype call (audio and visual)	1 (6)
No preference	8 (47)

^aChildren aged 9-23 years.

^bA total of 2 parents only reported Hispanic ethnicity and no racial category.

^cIncludes human papillomavirus, Tdap (tetanus, diphtheria, and pertussis), meningitis, and influenza vaccination.

^dHPV: human papillomavirus.

^en=43 children; received at least one human papillomavirus vaccine dose.

^fData missing for 5 parents.

Influences of HPV Vaccination Decision-Making: Qualitative Findings

Three themes emerged regarding parents' HPV vaccination decision-making processes: (1) barriers to HPV vaccination, (2) facilitators of HPV vaccination, and (3) suggestions for improving the HPV vaccination clinic experience.

Barriers to HPV Vaccination

Barriers that affected parents' HPV vaccination decision-making were driven by HPV misinformation and confusion, negative HPV beliefs and attitudes, and navigating trustworthy HPV information on the internet. Parents were unsure if boys needed or were eligible for the HPV vaccine. There was also concern that the vaccine did not cover all the "mutations of the virus" and that the "virus changed structure over time":

The HPV vaccine does not cover enough (too many types) for me to inject my son with some unknown drug. I feel like they are trying to scare us into doing it. I am not sold. [Female, 45 years, Black, non-Hispanic, very vaccine hesitant] (1)

A parent who had chosen not to vaccinate her son for HPV expressed the hope that people would get tested for STIs (including HPV) before having sex, although there is no Food and Drug Administration–approved HPV test for males. Another parent thought HPV was linked to herpes (although HPV does not cause herpes, it can cause genital warts). Parents questioned why "the vaccine series is not effective once you pass a certain age" and "does it work for those who are older than 26 but not sexually active?"

Parents reported negative beliefs and attitudes toward the HPV vaccine. Many parents who considered themselves generally pro-vaccination were skeptical about the HPV vaccine as they perceived it to be new and possibly have long-term negative health effects:

We are not getting it [HPV vaccine] right now. Maybe in a few years, when my daughter is closer to becoming sexually active and that way, after a few

more years have gone by, the vaccine will be more tested and better. [Female, 43 years, White, non-Hispanic, not vaccine hesitant] (2)

The belief of being able to postpone the HPV vaccine because of the perceived sexual inactivity of their adolescent was echoed by other parents:

I am reluctant to start the HPV shots with my soon to be 13-year old son because I don't think it's necessary at this moment to start it with him and I need to do more research to learn more about the risks of doing it or not doing it. [Female, 40 years, White, non-Hispanic, somewhat vaccine hesitant] (3)

I have always given my daughter the expected and recommended vaccinations except the HPV vaccination because of concerns I heard about negative effects, and also knowing absolutely that she is not sexually active yet I felt I had time to do some research and make an informed decision. [Female, 50 years, White, non-Hispanic, somewhat vaccine hesitant] (4)

Parents also questioned if the vaccine was being promoted for financial gain:

I just wonder if the doctors are urging these vaccines because it is what they believe is best or is big pharm pushing it. [Female, 45 years, Black, non-Hispanic, very vaccine hesitant] (5)

Parents generally felt the need to do their own HPV research, even when provided with recommendations by health care providers:

My pediatrician brought it [HPV vaccination] up at her last year's checkup (when she turned 12). She gave me a handout about the [HPV] vaccine and accepted my response that I was not ready to commit to it yet but wanted to research more and think about it. She did say she recommended it, but did not push, at least not at that visit. [Female, 50 years, White, non-Hispanic, somewhat vaccine hesitant] (6)

When doing their own HPV research, trustworthy information sources differed among parents. Although pharmaceutical advertisements and beliefs of the general public were seen as untrustworthy sources of HPV information and pediatricians were overwhelmingly seen as trustworthy HPV sources, there was little agreement among other sources. For example, family and friends, the internet, and news articles were HPV information sources that parents had differing opinions about:

TV news, old fashioned newspapers, and NPR are my primary sources for information. Generally they touch on a vaccine when it is new, recalled, or if there is a bump in a disease spreading. Not social media, not friends, not internet. I do read all the fine print in printed RX ads. [Female, 47 years, White, non-Hispanic, not vaccine hesitant] (7)

CDC website, friends and family, other news articles from magazines or website that are informative. I don't take my info from add in tv or magazine since they are paid by the pharmaceutical industry... [Female, 45 years, White, non-Hispanic, somewhat vaccine hesitant] (8)

Some parents noted the CDC and National Institutes of Health as trustworthy HPV information sources, whereas others commented that they had never thought of using these agencies:

...for some reason, I never thought to check the cdc about hpv, but after this focus group that makes a lot of sense...thank you all who have mentioned that [Female, 37 years, Black, non-Hispanic, vaccine hesitant] (9)

A parent also expressed knowing a CDC scientist who was not in support of the HPV vaccine. Navigating reputable research on the internet proved particularly challenging for many of the parents. This was highlighted in an exchange during a focus group where one parent shared a link to an article written by a physician disputing HPV vaccine effectiveness that *freaked her out*, and another parent provided a website citing that physician and his organization as fraudulent and not scientifically sound [46].

Overwhelmingly, parents had misinformation about HPV and the HPV vaccine and had differing opinions about *go-to* sources for reputable information. The HPV vaccine was seen as different from other adolescent vaccinations, even among those that considered themselves provaccine, because of the perceived newness, perceived lack of evidence, and belief that the vaccine could be postponed until a child is sexually active.

Facilitators of HPV Vaccination

Facilitators of HPV vaccination included positive HPV beliefs and attitudes, family members and close friends experiencing negative health outcomes from preventable illness, personal experience with HPV, and pediatrician recommendations. Positive beliefs about the HPV vaccine centered around decreasing their children's risk of acquiring STIs and cancer, protecting their children's future partners, and protecting public health. These beliefs were highlighted when parents were asked what comes to mind when they hear *HPV vaccine*, and statements around *cancer and sexually transmitted disease*

prevention were the most prominent. Parents expressed that the HPV vaccine provided a unique opportunity to prevent cancer:

Thank God for Modern Medicine. With technology and medical advancements, why are we not excited about the ability to PREVENT cancer- not just CURE it? Let's educate our parents and help them advocate for their children- it's time to be proactive now, rather than reactive later. [Female, 43 years, White, non-Hispanic, not vaccine hesitant] (10)

Most parents expressed that their child could be at risk of contracting HPV in the future. Parental perceived risk often centered on future partners:

As I said before, my hope is that my children choose to abstain from sex until marriage. Should they choose not to do that, I hope they use safe sex practices. But even if they did all the "right things," they still might wind up married to someone who was exposed to HPV. [Female, 37 years, White, non-Hispanic, not vaccine hesitant] (11)

Absolutely. Unfortunately, I think STD are very widespread. And I don't think people are generally forthcoming with telling others that they have an STD before having sex. [Female, 40 years, White, non-Hispanic, somewhat vaccine hesitant] (12)

Concerns over asymptomatic HPV infection and rape were also included in the discussion about risk. Parents mentioned that unlike other vaccines, HPV made them confront the impending adulthood of their adolescents:

...this is a vaccine that comes with a recognition that your child will someday be an adult. [Female, 44 years, White, non-Hispanic, somewhat vaccine hesitant] (13)

There is far more negative press associated with HPV and associating it with sexually transmitted disease. No one wants to think of their prepubescent child that way. It needs to be seen as a positive advancement in pediatric medicine. [Female, 43 years, White, non-Hispanic, not vaccine hesitant] (14)

Having a family member or close friend experience a negative health outcome related to HPV or another vaccine-preventable illness significantly influenced the way parents felt about the HPV vaccine:

I did not want to vaccinate my sons for HPV (even though it might be the best thing for public health) until I learned of a brother of a good friend who is gravely ill from HPV related throat cancer. [Female, 52 years, White, non-Hispanic, somewhat vaccine hesitant] (15)

One parent had HPV-attributed oropharyngeal cancer and became a vocal advocate for the HPV vaccine:

I know my initial rejection of the vaccination for my oldest son was that I was not raising him to be promiscuous, so I figured there was very little chance of him being exposed. Once I read the incredibly high transmission rates for the HPV virus and how easily

it is spread, and got to experience firsthand what the treatment was like, my oldest son was at the doctor receiving the vaccine immediately. I personally try to educate every parent I meet with younger children if the conversation gets steered to that subject. Usually all it takes is showing them my 5 inch scar from my neck dissection and relating how my taste buds are permanently destroyed from the radiation. [Male, 45 years, White, non-Hispanic, not vaccine hesitant] (16)

Overwhelmingly, parents expressed that a recommendation from their child's pediatrician positively influenced their decision to give their child the HPV vaccine, particularly if they were able to do their own HPV research before the clinic visit:

I was not sure about HPV for boys (and due to the fact that it's pretty new) but decided to follow recommendations from pediatrician and also from several articles that I read. [Female, 45 years, White, non-Hispanic, somewhat vaccine hesitant] (17)

I just recently became aware of HPV for boys from tv commercials and looked into it before my son was of age to receive it. At his last annual check-up, our pediatrician recommended the vaccine and I trusted her judgement enough to agree. [Female, 39 years, Black, non-Hispanic, not vaccine hesitant] (18)

There were additional factors that acted as both a barrier and facilitator to HPV vaccination. For example, adolescents influenced their parents' HPV vaccination decision-making process and were sometimes given autonomy over the decision:

My son heard a commercial about HPV and he wants to get the vaccine. My mom recently passed away from cancer and so he has a serious concern. [Female, 37 years, Black, non-Hispanic, vaccine hesitant] (19)

My second daughter chose not to receive the vaccine after discussing with our pediatrician. She is extremely mature and I was not going to insist she do something she felt strongly against. [Female, 52 years, White, non-Hispanic, somewhat vaccine hesitant] (20)

Parents expressed numerous positive beliefs and attitudes that encouraged them to get the HPV vaccine for their adolescent children. They told stories about how family, friends, and personal experiences with HPV shaped how they thought about HPV infection and outcomes. Pediatrician recommendation was one of the strongest facilitators for influencing a parent's decision-making.

Suggestions for Enhancing the HPV Vaccination Clinic Experience

Parents had ideas about enhancing communication with pediatricians and the clinic network that would help them feel more comfortable and informed about the HPV vaccine. Parents wanted HPV information and their pediatrician's HPV vaccination recommendation for their child to be sent to them months before their child's appointment:

It would be much preferred if [the clinic] was able to send information to parents 4-6 months before each round of any vaccine is due.... facts (amazing how little any of us in this group can site facts for the HPV!) along with your personal pediatrician's recommendation for your specific child: yay, nay, wait. Then it would be even better to have a chance to email or talk with the physician in advance of the appointment to ask questions. The HPV vaccine is a particularly awkward one to ask questions about in front of your 10 or 11 year old in the exam room. [Female, 47 years, White, non-Hispanic, not vaccine hesitant] (21)

Information in advance of appt would be helpful because if I hadn't already had experience with HPV vaccine I would have felt put on the spot when pedi asked if I wanted to give to him. [Female, 39 years, Black, non-Hispanic, not vaccine hesitant] (22)

It was also suggested that starting the conversation when the child is 9 years old would be helpful, along with the use of other clinic visits, such as the annual influenza vaccine appointment, to hand out informational materials. Parents wanted to be able to have a private dialog with the pediatrician if needed:

Provide info prior to the appointment, so parents can read and ask questions before going to the appointment. We don't really want to talk about risks, consequences, etc. in front of the child. [Female, 45 years old, White, non-Hispanic, somewhat vaccine hesitant] (23)

Begin HPV discussion with patient family at 9 year well visit. Bring it up again at 10 year well visit. Make it clear that questions are welcome, as well as private dialog via phone, if that would help parent. There is never enough time in the exam room to discuss things. [Female, 47 years, White, non-Hispanic, not vaccine hesitant] (24)

There was a spectrum of content and time spent by pediatricians in discussing the HPV vaccine with parents. Some parents had ongoing conversations with the pediatrician, whereas others knew very little:

We were told on several visits about the seriousness of the virus and the potential to cause cancer. [Male, 45 years, White, non-Hispanic, not vaccine hesitant] (25)

I have very little information about the vaccine. The doctor briefly mentioned it and said it's a good idea to do the series of three shots between the ages of 11-13. [Female, 40 years, White, non-Hispanic, somewhat vaccine hesitant] (26)

Parents reported a range of information they wanted from the pediatrician, everything from basic information to longitudinal research studies:

...a simple brochure with basic info and a guide to find other info would be nice to have [Female, 44 years, White, non-Hispanic, somewhat vaccine hesitant] (27)

I would like to see any studies that follow the patients that were the first recipients of the vaccine; I felt like the brochure that the pediatrician gave us to read was too dumbed down. I want real information, not a colorful brochure produced by the company that will profit from selling the vaccine. [Female, 52 years, White, non-Hispanic, somewhat vaccine hesitant] (28)

Parents suggested that partnering with school districts, having more information posted on the clinic network website, having an opt-in newsletter, and using social media to disseminate information would help them feel more informed about HPV. Parents were receptive to a clinic-sponsored HPV app if it included information they would want to receive from the pediatrician. Parents suggested including statistics on how widespread HPV is; how likely one is to be infected with it;

what the risks of being infected with HPV are, especially for boys; the benefits of receiving the vaccine; data on side effects and adverse reactions; how many years it has been available; and how thoroughly tested the HPV vaccine is compared with other vaccines.

Translating Findings Into DBCI Content and Feature Recommendations

After the focus group themes were identified, a wider research team, including pediatricians, behavioral scientists, statisticians, and designers, conducted multiple brainstorming sessions to translate the findings into DBCI content and feature recommendations focused on supporting the clinic network parents with their HPV vaccination decision-making processes. A total of 11 content and feature recommendations were suggested from qualitative themes (Table 2).

Table 2. Digital behavior change intervention (DBCI) content and feature design recommendations.

Recommendation	Description	Quote number ^b
Address HPV ^a knowledge barriers	Address prominent HPV and HPV vaccine knowledge barriers (ie, child being too young or sexually inexperienced, boys not being eligible, safety and side effects, and effectiveness)	1, 2, 3, and 4
Use trusted sources to educate and correct misinformation	Use pediatricians to communicate information as they are trusted and respected sources for children's health	17 and 18
Focus on HPV messaging that resonates with parents	Frame HPV information in a way that resonates with parents (ie, preventing cancer)	10, 11, 12, 13, and 14
Guide parents on navigating reputable HPV resources	Provide reputable HPV resources to parents and guide them in using best practices for navigating consumer health information on the internet	3, 4, 6, 7, 8, and 9
Describe reputable HPV research in a comprehensible format	Interpret and describe HPV scientific research in a comprehensible format (ie, plain language at sixth-grade level and infographics)	28
Communicate who is sponsoring the DBCI	Communicate that trusted sources (ie, pediatric clinic network) are sponsoring the product	5
Design for self-tailoring	Design for the spectrum of parent information needs from reviews of basic information to reviews of scientific studies	27 and 28
Design for a family audience	Design for engagement between family members, including adolescents who may influence their parent's decision-making	19 and 20
Design for reflection	Give parents the opportunity to reflect on the health experiences of others in their personal and extended networks to increase salience and relevancy	15 and 16
Organize and prepare for the clinic visit	Prepare parents for their child's clinic visit by having them organize their questions and concerns beforehand	21, 22, 23, and 24
Extend the clinic visit and enhance the clinic network	Create infrastructure that extends the clinic visit and leverages the clinic network so parents can better connect with others and needed information before and after the clinic visit	21, 22, 23, 24, 25, and 26

^aHPV: human papillomavirus.

^bThe numbers refer to the numbered quotations in the paper.

Content Strategy Recommendations

The research team made six content strategy recommendations. First, the DBCI content should address the prominent HPV knowledge barriers expressed by the parents. These knowledge barriers stemmed from misinformation or a lack of information and included a child being too young or sexually inexperienced for the HPV vaccine, boys not needing or not being eligible for the HPV vaccine, the HPV vaccine having side effects that make it unsafe, and the HPV vaccine not being effective. The second content strategy recommendation was to increase the influence

of the DBCI by having the network pediatricians be the voice that educates and communicates information, as they are highly trusted sources for obtaining quality advice about a child's health. This might be accomplished through role modeling via video recordings or a frequently asked questions section that features photos and interviews with the pediatricians. The third content strategy recommendation was to frame information in a way that resonates with the parents, such as messaging, which focuses on cancer prevention and decreasing future risk. The fourth recommendation was to help parents navigate reputable

HPV resources, as many parents felt it necessary to do their own HPV research before having their child vaccinated. This could involve educating on best practices for navigating consumer health information on the internet. The fifth recommendation was to help parents interact with HPV scientific research in a comprehensible format. A comprehensible format may involve featuring plain language interpretations and infographics of results from scientific studies. Finally, it was recommended to communicate who was sponsoring the app. The parents were skeptical of *big pharmaceutical companies* pushing anything related to the HPV vaccine, so making it clear that the app comes from trusted sources (ie, pediatric clinic network) may help improve its acceptability among parents.

Feature Design Recommendations

Five feature recommendations were made by the research team. The first was to design the DBCI for self-tailoring to meet the needs of individual parents. The HPV content parents wished to access differed in topic and depth as some parents were satisfied with the overview material, whereas others wanted to see scientific studies. Providing a self-tailored experience may be accomplished by presenting overview material but also offering links to reputable articles from the National Cancer Institute and CDC, giving parents the opportunity to explore a topic further if desired. The second recommendation was to include features that could engage a family audience, such as interactive games, as children's attitudes toward the vaccine could influence their parents' decision-making. The third recommendation was to have a feature where parents could reflect on the health experiences of others in their personal and extended networks. This might be advantageous for helping parents address anticipated regret and explore a more nuanced understanding of HPV. This might be accomplished by providing a guided prompt that parents can complete independently or with loved ones. As parents expressed limited time with the pediatricians during clinic visits, the fourth recommendation was to have a feature that organizes and records parent questions and pediatrician responses during the clinic visit. This may help increase the volume of information that can be discussed and reviewed. Finally, creating a DBCI that extends the clinic visit and enhances the clinic network infrastructure could address parents' desire to have more interaction and communication beyond the clinic visit. For example, these features might include having personalized adolescent vaccination recommendations sent before a child's appointment, having an opportunity to communicate about sensitive topics without the child being present, and connecting and learning from other parents in the clinic network.

Feasibility of Skype Synchronous Text-Based Focus Groups

The synchronous text-based focus groups were effective in gathering insights from parents with adolescents belonging to a geographically diverse pediatric clinic network. Most parents

were experienced using Skype *sometimes* (18/22, 82%). Parents requested a chat (text-only) format (8/17, 47%) over a call (audiovisual) format (1/17, 6%) for their focus group (Table 1). All parents successfully logged into their Skype accounts approximately 48 hours before their session and answered the welcome message as instructed. On the day of the session, all parents participated in their specified session. A total of 21 parents logged in and were ready to start on time (responded *yes* as instructed after reading the moderator's introductory message). One parent was 4 minutes late to their session but was able to respond to the first topic posted in the chat. During one of the sessions, two accidental group calls were made a few minutes apart by a parent. However, the accidental calls did not disconnect any parents and did not cause any disruption beyond confusion for a few seconds. No calls were made to the technology facilitator. Cost savings were US \$260 for each web-based session compared with the estimated costs for in-person sessions.

Key reflections on the feasibility and logistics of the format are highlighted in Table 3. Owing to the synchronous format, the sessions moved very quickly, with parents responding simultaneously at times. Preparing potential probes beforehand that could be easily modified and copy-pasted into the Skype chat was advantageous in keeping pace with participant responses. The format proved equitable with parents able to contribute to all relevant questions, and the Skype text bubbles, which occur when someone is typing, cued the study team into the cadence of posting questions. The format also produced automatic transcripts, allowing for immediate data analysis. As found in other studies, the anonymity provided by the format supported sensitive experiences being shared candidly [29,33]. For example, a parent discussed being sexually assaulted as an example of how rape might be a risk factor for acquiring HPV, and another discussed his own HPV cancer treatment. Web-based focus groups may be well-suited for discussions of HPV because of its association with sex, STIs, and reproductive cancers. An unexpected outcome of the web-based format was that parents shared articles and video clips in real time. It was beneficial for the study team to see examples of existing content that was influential in parents' decision to vaccinate. The format increased the utility of parents to succinctly exemplify media, enabling more efficient understanding by the study team and easier translation to inform intervention content recommendations. At the conclusion of the focus groups, the study team sent an email to all participants clarifying HPV misinformation, answering any HPV questions they posed during the focus groups and directing them to reputable sources. The text-based format had several key weaknesses, including insights from facial expression, physical gesturing and prosody being difficult to gather and discern, and the format being labor intensive, with multiple team members needed to help moderate the high demands of processing simultaneous information.

Table 3. Skype feasibility and logistics.

Category and description	Findings and experience	Reflections	
		Strength	Weakness
Format			
<ul style="list-style-type: none"> Participants were asked about their previous experience using Skype and were given the option to choose between a Skype chat-based focus group and a Skype audiovisual-based focus group. 	<ul style="list-style-type: none"> Most participants had experience with using Skype before their session. Participants preferred a chat-based format over an audiovisual format. The original study plan to have half the sessions be audiovisual and half be text-only for comparative reasons was abandoned when scheduling for the large percentage of participants that requested a text-based format became prohibitive. 	<ul style="list-style-type: none"> The automatic transcripts produced from the text-based format allowed for immediate qualitative data analysis supporting rapid formative research. 	<ul style="list-style-type: none"> Insights from facial expression, physical gesturing, and prosody were difficult to discern and gather with the text-based format.
Attendance			
<ul style="list-style-type: none"> Participants were instructed to log into their Skype study accounts 48 hours before their session and answer the welcome message. Participants were instructed to log-in a few minutes before their specified session and reply to the moderator's welcome instructions. 	<ul style="list-style-type: none"> All participants successfully responded to the welcome message. On the day of the session, all participants successfully responded to the welcome instructions and attended their specified session. Most attended on time. 	<ul style="list-style-type: none"> Participants were able to navigate the Skype chat function without issue. Attendance was high, possibly because of familiarity with the Skype platform and the ease of participating from a preferred location. 	<ul style="list-style-type: none"> It was difficult to verify the identities of participants.
Confidentiality and anonymity			
<ul style="list-style-type: none"> Skype usernames (parents' first name and ID number) and passwords were created for each participant and deleted at the conclusion of each session. 	<ul style="list-style-type: none"> By only using the participants' first names, they were able to recognize when someone was addressing them but still keep their identity anonymous. 	<ul style="list-style-type: none"> Sensitive experiences were shared candidly. 	<ul style="list-style-type: none"> It was difficult to verify the identities of participants.
Moderator considerations and cadence			
<ul style="list-style-type: none"> A total of 4 team members ran each session: the lead moderator, a tech facilitator, and 2 submoderators. 	<ul style="list-style-type: none"> At times, the session moved very quickly, with participants answering questions simultaneously. Skype chat provided text bubbles when a participant was typing which aided the research team in establishing the cadence of asking questions. Preparing potential probes that could be easily modified and copy-pasted into the Skype chat proved advantageous for keeping pace with participants. Having participants use Skype on a laptop or desktop with a connected keyboard rather than on their phone proved advantageous for more uniform and rapid response time. 	<ul style="list-style-type: none"> All participants were able to contribute to all relevant questions at their own speed. 	<ul style="list-style-type: none"> The type-based format took more team members to moderate than an in-person session because of the high demands of processing incoming information spurred by simultaneous typing and posting.
Group dynamics			

Category and description	Findings and experience	Reflections	
		Strength	Weakness
<ul style="list-style-type: none"> Unexpected outcomes occurred from the synchronous web-based format. 	<ul style="list-style-type: none"> Participants shared HPV^a articles and video clips that influenced their decision to vaccinate in real time. At the conclusion of the focus groups, the team sent an email to all participants clarifying HPV misinformation, answering any HPV questions they posed during the focus groups and directing them to reputable sources. 	<ul style="list-style-type: none"> It was beneficial for the research team to see examples of actual content that influenced participants' decision to vaccinate. The format increased the utility of participants to succinctly exemplify media, enabling more efficient understanding by the study team. 	<ul style="list-style-type: none"> It was important to clarify misinformation after the sessions as participants shared articles that used persuasive tactics of expert opinion and pseudoscience to discredit the HPV vaccine.
Tech support and disruption			
<ul style="list-style-type: none"> A 'how to download and use Skype' guide with screenshots was distributed before the sessions. A technology facilitator was available during the sessions for support. 	<ul style="list-style-type: none"> No assistance was needed downloading or operating Skype. Two brief accidental calls occurred during one session but did not cause major disruption or require intervention. Technology support was not actively needed during the sessions. 	<ul style="list-style-type: none"> Participant familiarity and experience with Skype supported efficient operations. 	<ul style="list-style-type: none"> Although not needed, it may have proved difficult to help participants navigate technical challenges remotely.
Cost savings			
<ul style="list-style-type: none"> The costs of conducting the focus groups web-based versus in person were compared. 	<ul style="list-style-type: none"> Cost savings were estimated at US \$260 for each session, accounting for transcription, parking reimbursements, and clinic staff after-hours pay. 	<ul style="list-style-type: none"> Skype is a free platform and provides cost savings compared with in-person methods. 	<ul style="list-style-type: none"> The format may bias participation toward those with access to a computer and internet and who feel comfortable with web-based communication.

^aHPV: human papillomavirus.

Discussion

Principal Findings

This study highlights the factors that influence HPV vaccine decision-making among a group of mostly White, non-Hispanic, and educated parents whose adolescents were patients at a large urban pediatric clinic network in the United States. The findings from these synchronous text-based focus groups align with findings found in other HPV vaccination studies conducted with parents [15,18,28,47,48]. Even among these highly educated individuals, where many expressed interactions with and access to health care professionals, there was confusion, misinformation, and a lack of knowledge regarding HPV and the HPV vaccine. This has been a consistent challenge since the introduction of the first HPV vaccine in 2006 [49,50]. Similar to other studies, parents generally had favorable attitudes toward adolescent vaccinations but differed in how they viewed the HPV vaccine [47]. Also consistent with prior findings is the role pediatrician recommendation serves in positively influencing HPV vaccination parental decision-making [15,48]. Personal or familial experience with HPV had a significant

influence on HPV vaccination attitudes and beliefs, even leading some parents to become outspoken advocates for the vaccine. Conducting their own research on the web was an important step in parent HPV vaccination decision-making and could lead to missed opportunities for vaccination if the parent had not yet researched HPV or was not prepared to discuss their outstanding questions with the pediatrician at the clinic visit. It is not surprising that the parents had a difficult time discerning reputable web-based research, as antivaccination websites often use expert opinion (using the title of *Dr*) and pseudoscientific evidence (confusing correlation for causation) as persuasive tactics [51].

These results indicate that parents want more communication from their child's pediatrician and have certain questions and advice they are actively seeking. Pediatric clinic networks have the opportunity to cultivate credible, relatable, and understandable HPV information for distribution to parents and patients before a clinic visit to increase HPV vaccination rates. Content and feature recommendations from this study can provide guidance for researchers and developers involved in the creation of HPV-focused apps for parents. These recommendations are largely consistent with general

recommendations for digital health interventions and are related primarily to mitigating concerns and misinformation, providing authentic and persuasive messages, providing user control in inquiries of content breadth and depth, and facilitating a move to immediate action (vaccination appointments) [52,53]. In addition, features and content that incorporate personalization, reinforcement learning, social support, credibility of sources, and focus on simple and consistent interface esthetics, easy-to-use navigation, and multimedia messages have been found to influence and improve user participation across health topics [54]. The recommendations emerging from this work provide insight but should not be regarded as definitive. Future research and development that uses well-validated theory and empirically-based development frameworks to design and formatively evaluate proofs-of-concept and prototypes is recommended [22-25]. In this regard, this study represents an important initial needs assessment step to give voice to the patient and parent experience and translate it into a responsive intervention.

Recent events of the COVID-19 pandemic have seen a dramatic movement toward web-based communication across transaction domains, including research protocols. Focus groups are well accepted as a needs assessment method, and now platforms such as Zoom and WebEx are ubiquitous and offer multiple features, including video, chat, questions and answers, and break-out room features to facilitate them. This study presents results that can provide useful insights when considering the efficacy of synchronous text-based focus groups and the advantages of this method. The preference of the participants to opt for text-based communication was unexpected; however, it was consistent with a desire for confidentiality, anonymity, convenience, and avoidance of video-related bandwidth and other technical problems. Furthermore, the text-based format provided response equity, enabling each participant to submit responses irrespective of their comfort in group situations. Parents participating in these focus groups approximated the demographics of the clinic

network population as predominately White, non-Hispanic, and privately insured. However, a limitation of the study is that the web-based format may have excluded participation from parents with lower socioeconomic status or those who did not feel comfortable with web-based communication. The results may not be generalizable to parents who are younger, have lower educational attainment, who are publicly insured, from rural communities, or are not White and non-Hispanic. Future research addressing these populations is recommended. Although seemingly amenable to information gathering, this study was exploratory and did not determine personal determinants that might aid recruitment into such groups or potentially bias information gained in this type of forum. Further concerns about this approach are the impact of lack of nonverbal feedback from the group and the potential challenge of ensuring accountability of participants in focusing on the discussion. Further research in this regard is recommended. The method did offer significant parsimony from a logistic perspective, as the text-based approach had distinct advantages for streamlining data processing, management, and analysis when compared with audio or video recordings.

Conclusions

Synchronous text-based focus groups conducted via Skype are feasible for conducting DBCI formative research on HPV vaccination decision-making. Among this well-educated and well-resourced parent sample, there were barriers such as misinformation and facilitators such as pediatrician recommendations that influenced HPV vaccination decision-making. Parents want easy-to-understand and relevant HPV vaccination advice from their child's pediatrician and an enhanced clinic visit experience, which lets them access and connect to tailored information before and after clinic visits. The results gathered provide guidance for content and features that may inform a more responsive DBCI to address HPV vaccination decision-making among parents.

Acknowledgments

The authors thank the parents who participated in the study for their time and valuable insights. The authors also thank doctoral student, Preena Loomba, for her help in conducting the focus groups. This research was supported by the Cancer Prevention Institute of Texas RP150014.

Conflicts of Interest

None declared.

References

1. Satterwhite C, Torrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MC, et al. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2008. *Sex Transm Dis* 2013 Mar;40(3):187-193. [doi: [10.1097/OLQ.0b013e318286bb53](https://doi.org/10.1097/OLQ.0b013e318286bb53)] [Medline: [23403598](https://pubmed.ncbi.nlm.nih.gov/23403598/)]
2. Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, et al. Global burden of human papillomavirus and related diseases. *Vaccine* 2012 Nov 20;30 Suppl 5:12-23. [doi: [10.1016/j.vaccine.2012.07.055](https://doi.org/10.1016/j.vaccine.2012.07.055)] [Medline: [23199955](https://pubmed.ncbi.nlm.nih.gov/23199955/)]
3. Chesson HW, Dunne EF, Hariri S, Markowitz LE. The estimated lifetime probability of acquiring human papillomavirus in the United States. *Sex Transm Dis* 2014 Nov;41(11):660-664 [FREE Full text] [doi: [10.1097/OLQ.000000000000193](https://doi.org/10.1097/OLQ.000000000000193)] [Medline: [25299412](https://pubmed.ncbi.nlm.nih.gov/25299412/)]

4. Moscicki AB, Hills N, Shiboski S, Powell K, Jay N, Hanson E, et al. Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females. *J Am Med Assoc* 2001 Jun 20;285(23):2995-3002. [doi: [10.1001/jama.285.23.2995](https://doi.org/10.1001/jama.285.23.2995)] [Medline: [11410098](https://pubmed.ncbi.nlm.nih.gov/11410098/)]
5. Winer R, Lee SK, Hughes JP, Adam DE, Kiviat NB, Koutsky LA. Genital human papillomavirus infection: incidence and risk factors in a cohort of female university students. *Am J Epidemiol* 2003 Mar 01;157(3):218-226. [doi: [10.1093/aje/kwf180](https://doi.org/10.1093/aje/kwf180)] [Medline: [12543621](https://pubmed.ncbi.nlm.nih.gov/12543621/)]
6. Dunne EF, Markowitz LE. Genital human papillomavirus infection. *Clin Infect Dis* 2006 Sep 01;43(5):624-629. [doi: [10.1086/505982](https://doi.org/10.1086/505982)] [Medline: [16886157](https://pubmed.ncbi.nlm.nih.gov/16886157/)]
7. Senkomago V, Henley SJ, Thomas CC, Mix JM, Markowitz LE, Saraiya M. Human papillomavirus-attributable cancers - United States, 2012-2016. *MMWR Morb Mortal Wkly Rep* 2019 Aug 23;68(33):724-728 [FREE Full text] [doi: [10.15585/mmwr.mm6833a3](https://doi.org/10.15585/mmwr.mm6833a3)] [Medline: [31437140](https://pubmed.ncbi.nlm.nih.gov/31437140/)]
8. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol* 2011 Nov 10;29(32):4294-4301 [FREE Full text] [doi: [10.1200/JCO.2011.36.4596](https://doi.org/10.1200/JCO.2011.36.4596)] [Medline: [21969503](https://pubmed.ncbi.nlm.nih.gov/21969503/)]
9. Hoy T, Singhal PK, Willey VJ, Insinga RP. Assessing incidence and economic burden of genital warts with data from a US commercially insured population. *Curr Med Res Opin* 2009 Oct;25(10):2343-2351. [doi: [10.1185/03007990903136378](https://doi.org/10.1185/03007990903136378)] [Medline: [19650749](https://pubmed.ncbi.nlm.nih.gov/19650749/)]
10. Markowitz LE, Dunne EF, Saraiya M, Chesson HW, Curtis CR, Gee J, Centers for Disease Control and Prevention (CDC). Human papillomavirus vaccination: recommendations of the advisory committee on immunization practices (ACIP). *MMWR Recomm Rep* 2014 Aug 29;63(RR-05):1-30 [FREE Full text] [Medline: [25167164](https://pubmed.ncbi.nlm.nih.gov/25167164/)]
11. Petrosky E, Bocchini JA, Hariri S, Chesson H, Curtis CR, Saraiya M, et al. Use of 9-valent human papillomavirus (HPV) vaccine: updated HPV vaccination recommendations of the advisory committee on immunization practices. *MMWR Morb Mortal Wkly Rep* 2015 Mar 27;64(11):300-304 [FREE Full text] [Medline: [25811679](https://pubmed.ncbi.nlm.nih.gov/25811679/)]
12. Markowitz LE, Gee J, Chesson H, Stokley S. Ten years of human papillomavirus vaccination in the United States. *Acad Pediatr* 2018 Mar;18(2S):3-10. [doi: [10.1016/j.acap.2017.09.014](https://doi.org/10.1016/j.acap.2017.09.014)] [Medline: [29502635](https://pubmed.ncbi.nlm.nih.gov/29502635/)]
13. Meites E, Kempe A, Markowitz LE. Use of a 2-dose schedule for human papillomavirus vaccination - updated recommendations of the advisory committee on immunization practices. *MMWR Morb Mortal Wkly Rep* 2016 Dec 16;65(49):1405-1408 [FREE Full text] [doi: [10.15585/mmwr.mm6549a5](https://doi.org/10.15585/mmwr.mm6549a5)] [Medline: [27977643](https://pubmed.ncbi.nlm.nih.gov/27977643/)]
14. Walker TY, Elam-Evans LD, Singleton JA, Yankey D, Markowitz LE, Fredua B, et al. National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years - United States, 2016. *MMWR Morb Mortal Wkly Rep* 2017 Aug 25;66(33):874-882 [FREE Full text] [doi: [10.15585/mmwr.mm6633a2](https://doi.org/10.15585/mmwr.mm6633a2)] [Medline: [28837546](https://pubmed.ncbi.nlm.nih.gov/28837546/)]
15. Rodriguez SA, Mullen PD, Lopez DM, Savas LS, Fernández ME. Factors associated with adolescent HPV vaccination in the U.S.: a systematic review of reviews and multilevel framework to inform intervention development. *Prev Med* 2020 Feb;131:105968 [FREE Full text] [doi: [10.1016/j.ypmed.2019.105968](https://doi.org/10.1016/j.ypmed.2019.105968)] [Medline: [31881235](https://pubmed.ncbi.nlm.nih.gov/31881235/)]
16. Head KJ, Biederman E, Sturm LA, Zimet GD. A retrospective and prospective look at strategies to increase adolescent HPV vaccine uptake in the United States. *Hum Vaccin Immunother* 2018 Jul 03;14(7):1626-1635 [FREE Full text] [doi: [10.1080/21645515.2018.1430539](https://doi.org/10.1080/21645515.2018.1430539)] [Medline: [29359986](https://pubmed.ncbi.nlm.nih.gov/29359986/)]
17. Gidengil C, Chen C, Parker AM, Nowak S, Matthews L. Beliefs around childhood vaccines in the United States: a systematic review. *Vaccine* 2019 Oct 23;37(45):6793-6802 [FREE Full text] [doi: [10.1016/j.vaccine.2019.08.068](https://doi.org/10.1016/j.vaccine.2019.08.068)] [Medline: [31562000](https://pubmed.ncbi.nlm.nih.gov/31562000/)]
18. Beavis A, Krakow M, Levinson K, Rositch AF. Reasons for lack of HPV vaccine initiation in NIS-teen over time: shifting the focus from gender and sexuality to necessity and safety. *J Adolesc Health* 2018 Nov;63(5):652-656. [doi: [10.1016/j.jadohealth.2018.06.024](https://doi.org/10.1016/j.jadohealth.2018.06.024)] [Medline: [30348283](https://pubmed.ncbi.nlm.nih.gov/30348283/)]
19. Hughes J, Cates JR, Liddon N, Smith JS, Gottlieb SL, Brewer NT. Disparities in how parents are learning about the human papillomavirus vaccine. *Cancer Epidemiol Biomarkers Prev* 2009 Mar;18(2):363-372 [FREE Full text] [doi: [10.1158/1055-9965.EPI-08-0418](https://doi.org/10.1158/1055-9965.EPI-08-0418)] [Medline: [19190161](https://pubmed.ncbi.nlm.nih.gov/19190161/)]
20. MacLean SA, Basch CH, Ethan D, Garcia P. Readability of online information about HPV Immunization. *Hum Vaccin Immunother* 2019;15(7-8):1505-1507 [FREE Full text] [doi: [10.1080/21645515.2018.1502518](https://doi.org/10.1080/21645515.2018.1502518)] [Medline: [30024828](https://pubmed.ncbi.nlm.nih.gov/30024828/)]
21. Phillips K. No, bananas don't cure HIV, nor will garlic cure COVID-19: searching for, assessing, and consuming health information online. *J Consum Health Internet* 2020 Jun 03;24(2):175-185. [doi: [10.1080/15398285.2020.1755149](https://doi.org/10.1080/15398285.2020.1755149)]
22. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res* 2015 Jan 30;17(1):e30 [FREE Full text] [doi: [10.2196/jmir.4055](https://doi.org/10.2196/jmir.4055)] [Medline: [25639757](https://pubmed.ncbi.nlm.nih.gov/25639757/)]
23. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (integrate, design, assess, and share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. *J Med Internet Res* 2016 Dec 16;18(12):e317 [FREE Full text] [doi: [10.2196/jmir.5927](https://doi.org/10.2196/jmir.5927)] [Medline: [27986647](https://pubmed.ncbi.nlm.nih.gov/27986647/)]
24. Patel S, Arya M. The BUS framework: a comprehensive tool in creating an mHealth app utilizing behavior change theories, user-centered design, and social marketing. *J Mob Technol Med* 2017 Apr;6(1):39-45 [FREE Full text] [doi: [10.7309/jmtm.6.1.6](https://doi.org/10.7309/jmtm.6.1.6)] [Medline: [29104674](https://pubmed.ncbi.nlm.nih.gov/29104674/)]

25. Sucala M, Ezeanochie N, Cole-Lewis H, Turgiss J. An iterative, interdisciplinary, collaborative framework for developing and evaluating digital behavior change interventions. *Transl Behav Med* 2020 Dec 31;10(6):1538-1548 [FREE Full text] [doi: [10.1093/tbm/ibz109](https://doi.org/10.1093/tbm/ibz109)] [Medline: [31328775](https://pubmed.ncbi.nlm.nih.gov/31328775/)]
26. Stiles-Shields C, Plevinsky J, Psihogios A, Holmbeck G. Considerations and future directions for conducting clinical research with pediatric populations during the COVID-19 pandemic. *J Pediatr Psychol* 2020 Jul 17;45(7):720-724. [doi: [10.1093/jpepsy/jsaa055](https://doi.org/10.1093/jpepsy/jsaa055)]
27. Saberi P. Research in the time of coronavirus: continuing ongoing studies in the midst of the COVID-19 pandemic. *AIDS Behav* 2020 Aug;24(8):2232-2235 [FREE Full text] [doi: [10.1007/s10461-020-02868-4](https://doi.org/10.1007/s10461-020-02868-4)] [Medline: [32303924](https://pubmed.ncbi.nlm.nih.gov/32303924/)]
28. Fontenot HB, Domush V, Zimet GD. Parental attitudes and beliefs regarding the nine-valent human papillomavirus vaccine. *J Adolesc Health* 2015 Dec;57(6):595-600. [doi: [10.1016/j.jadohealth.2015.09.003](https://doi.org/10.1016/j.jadohealth.2015.09.003)] [Medline: [26592327](https://pubmed.ncbi.nlm.nih.gov/26592327/)]
29. Fontenot HB, Rosenberger JG, McNair KT, Mayer KH, Zimet G. Perspectives and preferences for a mobile health tool designed to facilitate HPV vaccination among young men who have sex with men. *Hum Vaccin Immunother* 2019;15(7-8):1815-1823 [FREE Full text] [doi: [10.1080/21645515.2019.1568156](https://doi.org/10.1080/21645515.2019.1568156)] [Medline: [30625049](https://pubmed.ncbi.nlm.nih.gov/30625049/)]
30. Wettergren L, Eriksson LE, Nilsson J, Jervaeus A, Lampic C. Online focus group discussion is a valid and feasible mode when investigating sensitive topics among young persons with a cancer experience. *JMIR Res Protoc* 2016 May 09;5(2):e86 [FREE Full text] [doi: [10.2196/resprot.5616](https://doi.org/10.2196/resprot.5616)] [Medline: [27161146](https://pubmed.ncbi.nlm.nih.gov/27161146/)]
31. Jervaeus A, Nilsson J, Eriksson LE, Lampic C, Widmark C, Wettergren L. Exploring childhood cancer survivors' views about sex and sexual experiences -findings from online focus group discussions. *Eur J Oncol Nurs* 2016 Mar;20:165-172 [FREE Full text] [doi: [10.1016/j.ejon.2015.07.009](https://doi.org/10.1016/j.ejon.2015.07.009)] [Medline: [26251362](https://pubmed.ncbi.nlm.nih.gov/26251362/)]
32. Namey E, Guest G, O'Regan A, Godwin CL, Taylor J, Martinez A. How does mode of qualitative data collection affect data and cost? Findings from a quasi-experimental study. *Field Methods* 2019 Dec 16;32(1):58-74. [doi: [10.1177/1525822x19886839](https://doi.org/10.1177/1525822x19886839)]
33. Woodyatt CR, Finneran CA, Stephenson R. In-person versus online focus group discussions: a comparative analysis of data quality. *Qual Health Res* 2016 May;26(6):741-749. [doi: [10.1177/1049732316631510](https://doi.org/10.1177/1049732316631510)] [Medline: [26935719](https://pubmed.ncbi.nlm.nih.gov/26935719/)]
34. Campbell MK, Meier A, Carr C, Enga Z, James AS, Reedy J, et al. Health behavior changes after colon cancer: a comparison of findings from face-to-face and on-line focus groups. *Fam Community Health* 2001 Oct;24(3):88-103. [doi: [10.1097/00003727-200110000-00010](https://doi.org/10.1097/00003727-200110000-00010)] [Medline: [11563947](https://pubmed.ncbi.nlm.nih.gov/11563947/)]
35. Skype. Microsoft. URL: <https://www.skype.com/en/> [accessed 2021-09-30]
36. Opel DJ, Mangione-Smith R, Taylor JA, Korfiatis C, Wiese C, Catz S, et al. Development of a survey to identify vaccine-hesitant parents: the parent attitudes about childhood vaccines survey. *Hum Vaccin* 2011 Apr;7(4):419-425 [FREE Full text] [doi: [10.4161/hv.7.4.14120](https://doi.org/10.4161/hv.7.4.14120)] [Medline: [21389777](https://pubmed.ncbi.nlm.nih.gov/21389777/)]
37. Alexander AB, Stupiansky NW, Ott MA, Herbenick D, Reece M, Zimet GD. What parents and their adolescent sons suggest for male HPV vaccine messaging. *Health Psychol* 2014 May;33(5):448-456. [doi: [10.1037/a0033863](https://doi.org/10.1037/a0033863)] [Medline: [24588632](https://pubmed.ncbi.nlm.nih.gov/24588632/)]
38. Paskett ED, Krok-Schoen JL, Pennell ML, Tatum CM, Reiter PL, Peng J, et al. Results of a multilevel intervention trial to increase human papillomavirus (HPV) vaccine uptake among adolescent girls. *Cancer Epidemiol Biomarkers Prev* 2016 Apr;25(4):593-602 [FREE Full text] [doi: [10.1158/1055-9965.EPI-15-1243](https://doi.org/10.1158/1055-9965.EPI-15-1243)] [Medline: [27196093](https://pubmed.ncbi.nlm.nih.gov/27196093/)]
39. Fernández ME, Le YL, Fernández-Espada N, Calo WA, Savas LS, Vélez C, et al. Knowledge, attitudes, and beliefs about human papillomavirus (HPV) vaccination among Puerto Rican mothers and daughters, 2010: a qualitative study. *Prev Chronic Dis* 2014 Dec 04;11:E212 [FREE Full text] [doi: [10.5888/pcd11.140171](https://doi.org/10.5888/pcd11.140171)] [Medline: [25474384](https://pubmed.ncbi.nlm.nih.gov/25474384/)]
40. Bond SM, Cartmell KB, Lopez CM, Ford ME, Brandt HM, Gore EI, et al. Racial and ethnic group knowledge, perceptions and behaviors about human papillomavirus, human papillomavirus vaccination, and cervical cancer among adolescent females. *J Pediatr Adolesc Gynecol* 2016 Oct;29(5):429-435 [FREE Full text] [doi: [10.1016/j.jpog.2016.02.005](https://doi.org/10.1016/j.jpog.2016.02.005)] [Medline: [26915922](https://pubmed.ncbi.nlm.nih.gov/26915922/)]
41. Shafer A, Cates JR, Diehl SJ, Hartmann M. Asking mom: formative research for an HPV vaccine campaign targeting mothers of adolescent girls. *J Health Commun* 2011 Oct;16(9):988-1005. [doi: [10.1080/10810730.2011.571343](https://doi.org/10.1080/10810730.2011.571343)] [Medline: [21728780](https://pubmed.ncbi.nlm.nih.gov/21728780/)]
42. Barnack-Tavlaris JL, Garcini L, Sanchez O, Hernandez I, Navarro AM. Focus group discussions in community-based participatory research to inform the development of a human papillomavirus (HPV) educational intervention for Latinas in San Diego. *J Cancer Educ* 2013 Dec;28(4):784-789 [FREE Full text] [doi: [10.1007/s13187-013-0516-7](https://doi.org/10.1007/s13187-013-0516-7)] [Medline: [23857185](https://pubmed.ncbi.nlm.nih.gov/23857185/)]
43. Reiter PL, Oldach BR, Randle KE, Katz ML. Acceptability of HPV vaccine for males and preferences for future education programs among Appalachian residents. *Am J Mens Health* 2014 Mar;8(2):167-174 [FREE Full text] [doi: [10.1177/1557988313505319](https://doi.org/10.1177/1557988313505319)] [Medline: [24085197](https://pubmed.ncbi.nlm.nih.gov/24085197/)]
44. 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/)]
45. Creswell JW, Poth CN. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: SAGE Publications; 2016.
46. Barrett S. Dr. Joseph mercola ordered to stop illegal claims. Quackwatch. 2016. URL: <https://quackwatch.org/11Ind/mercola/> [accessed 2020-08-16]

47. Ogunbajo A, Hansen CE, North AL, Okoloko E, Niccolai LM. "I think they're all basically the same": parents' perceptions of human papilloma virus (HPV) vaccine compared with other adolescent vaccines. *Child Care Health Dev* 2016 Jul;42(4):582-587 [[FREE Full text](#)] [doi: [10.1111/cch.12331](https://doi.org/10.1111/cch.12331)] [Medline: [26990908](https://pubmed.ncbi.nlm.nih.gov/26990908/)]
48. Holman DM, Benard V, Roland KB, Watson M, Liddon N, Stokley S. Barriers to human papillomavirus vaccination among US adolescents: a systematic review of the literature. *JAMA Pediatr* 2014 Jan;168(1):76-82 [[FREE Full text](#)] [doi: [10.1001/jamapediatrics.2013.2752](https://doi.org/10.1001/jamapediatrics.2013.2752)] [Medline: [24276343](https://pubmed.ncbi.nlm.nih.gov/24276343/)]
49. Dorell C, Yankey D, Jeyarajah J, Stokley S, Fisher A, Markowitz L, et al. Delay and refusal of human papillomavirus vaccine for girls, national immunization survey-teen, 2010. *Clin Pediatr (Phila)* 2014 Mar;53(3):261-269 [[FREE Full text](#)] [doi: [10.1177/0009922813520070](https://doi.org/10.1177/0009922813520070)] [Medline: [24463951](https://pubmed.ncbi.nlm.nih.gov/24463951/)]
50. Dorell C, Yankey D, Strasser S. Parent-reported reasons for nonreceipt of recommended adolescent vaccinations, national immunization survey: teen, 2009. *Clin Pediatr (Phila)* 2011 Dec;50(12):1116-1124. [doi: [10.1177/0009922811415104](https://doi.org/10.1177/0009922811415104)] [Medline: [21856964](https://pubmed.ncbi.nlm.nih.gov/21856964/)]
51. Moran MB, Lucas M, Everhart K, Morgan A, Prickett E. What makes anti-vaccine websites persuasive? A content analysis of techniques used by anti-vaccine websites to engender anti-vaccine sentiment. *J Commun Healthc* 2016 Oct 03;9(3):151-163. [doi: [10.1080/17538068.2016.1235531](https://doi.org/10.1080/17538068.2016.1235531)]
52. Morrison LG. Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: a review. *Digit Health* 2015 Jul 17;1:2055207615595335 [[FREE Full text](#)] [doi: [10.1177/2055207615595335](https://doi.org/10.1177/2055207615595335)] [Medline: [29942544](https://pubmed.ncbi.nlm.nih.gov/29942544/)]
53. Wang Y, Fadhil A, Lange J, Reiterer H. Integrating taxonomies into theory-based digital health interventions for behavior change: a holistic framework. *JMIR Res Protoc* 2019 Jan 15;8(1):e8055 [[FREE Full text](#)] [doi: [10.2196/resprot.8055](https://doi.org/10.2196/resprot.8055)] [Medline: [30664477](https://pubmed.ncbi.nlm.nih.gov/30664477/)]
54. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. *J Med Internet Res* 2020 Dec 09;22(12):e21687 [[FREE Full text](#)] [doi: [10.2196/21687](https://doi.org/10.2196/21687)] [Medline: [33295292](https://pubmed.ncbi.nlm.nih.gov/33295292/)]

Abbreviations

CDC: US Centers for Disease Control and Prevention

DBCI: digital behavior change intervention

HPV: human papillomavirus

STI: sexually transmitted infection

Tdap: tetanus, diphtheria, and pertussis

Edited by G Eysenbach; submitted 25.04.21; peer-reviewed by G Sabben, A Morelli, Y Wei; comments to author 11.06.21; revised version received 25.07.21; accepted 18.09.21; published 19.11.21.

Please cite as:

Becker ERB, Shegog R, Savas LS, Frost EL, Healy CM, Spinner SW, Vernon SW

Informing Content and Feature Design of a Parent-Focused Human Papillomavirus Vaccination Digital Behavior Change Intervention: Synchronous Text-Based Focus Group Study

JMIR Form Res 2021;5(11):e28846

URL: <https://formative.jmir.org/2021/11/e28846>

doi: [10.2196/28846](https://doi.org/10.2196/28846)

PMID: [34806991](https://pubmed.ncbi.nlm.nih.gov/34806991/)

©Elisabeth RB Becker, Ross Shegog, Lara S Savas, Erica L Frost, C Mary Healy, Stanley W Spinner, Sally W Vernon. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 19.11.2021. 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

Guidance for Implementing Video Consultations in Danish General Practice: Rapid Cycle Coproduction Study

Amanda Sandbæk^{1,2*}, MSc; Line Due Christensen^{1*}, MSc, PhD; Lotte Lykke Larsen³, MSc; Nina Primholdt Christensen^{4,5}, MSc; Frida Greek Kofod⁶, MD; Ann Dorrit Guassora⁶, MSc, PhD; Camilla Hoffmann Merrild³, MSc, PhD; Elisabeth Assing Hvidt⁴, MSc, PhD

¹Research Unit for General Practice, Aarhus, Denmark

²Department of Public Health, Aarhus University, Aarhus, Denmark

³Center for General Practice at Aalborg University, Aalborg, Denmark

⁴Research Unit of General Practice, Department of Public Health, University of Southern Denmark, Odense, Denmark

⁵Department of Hematology, Odense University Hospital, Odense, Denmark

⁶The Research Unit and Section of General Practice, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

* these authors contributed equally

Corresponding Author:

Line Due Christensen, MSc, PhD
Research Unit for General Practice
Bartholins Allé 2
Aarhus, 8000
Denmark
Phone: 45 22791482
Email: linedue@ph.au.dk

Abstract

Background: The COVID-19 pandemic has changed various spheres of health care. General practitioners (GPs) have widely replaced face-to-face consultations with telephone or video consultations (VCs) to reduce the risk of COVID-19 transmission. Using VCs for health service delivery is an entirely new way of practicing for many GPs. However, this transition process has largely been conducted with no formal guidelines, which may have caused implementation barriers. This study presents a rapid cycle coproduction approach for developing a guide to assist VC implementation in general practice.

Objective: The aim of this paper is to describe the developmental phases of the VC guide to assist general practices in implementing VCs and summarize the evaluation made by general practice users.

Methods: The development of a guide for VC in general practice was structured as a stepped process based on the coproduction and prototyping processes. We used an iterative framework based on rapid qualitative analyses and interdisciplinary collaborations. Thus, the guide was developed in small, repeated cycles of development, implementation, evaluation, and adaptation, with a continuous exchange between research and practice. The data collection process was structured in 3 main phases. First, we conducted a literature review, recorded observations, and held informal and semistructured interviews. Second, we facilitated coproduction with stakeholders through 4 workshops with GPs, a group interview with patient representatives, and individual revisions by GPs. Third, nationwide testing was conducted in 5 general practice clinics and was followed by an evaluation of the guide through interviews with GPs.

Results: A rapid cycle coproduction approach was used to explore the needs of general practice in connection with the implementation of VC and to develop useful, relevant, and easily understandable guiding materials. Our findings suggest that a guide for VCs should include advice and recommendations regarding the organization of VCs, the technical setup, the appropriate target groups, patients' use of VCs, the performance of VCs, and the arrangements for booking a VC.

Conclusions: The combination of coproduction, prototyping, small iterations, and rapid data analysis is a suitable approach when contextually rich, hands-on guide materials are urgently needed. Moreover, this method could provide an efficient way of developing relevant guide materials for general practice to aid the implementation of new technology beyond the pandemic period.

(JMIR Form Res 2021;5(11):e27323) doi:[10.2196/27323](https://doi.org/10.2196/27323)

KEYWORDS

general practice; remote consultation; implementation science; resource guide; communication; video consultation; coproduction; rapid analysis; workshop; intervention development

Introduction

Background

When COVID-19 was declared a pandemic by the World Health Organization in March 2020, face-to-face consultations were largely replaced by telephone, email, or video consultations (VCs) in general practice in Denmark, as in many other Western countries, to reduce the risk of COVID-19 transmission [1]. These changes were recommended by the Danish Regions and the Danish Organization of General Practitioners [2].

Although the sudden implementation of VC was triggered by necessity, remote communication between patients and health care providers has been recommended as a sustainable solution for some health service deliveries to ensure increased access to health care [3,4]. VC has been known to save time and reduce transportation costs for patients, to be suitable for patients unable to visit the clinic, and to be advantageous for patients who are uncomfortable with visiting a clinic (eg, patients with certain mental health problems) [3,5-7]. However, VC had only been used sparingly in general practice in Denmark before the COVID-19 pandemic [8,9].

To enable VC in general practice, the Danish Organization of General Practitioners developed access to VC through the mobile app *My Doctor* at the beginning of the COVID-19 pandemic. The *My Doctor* app is a multifaceted tool for patient-provider communication. Besides facilitating VCs, the app includes the possibility of renewing prescriptions, keeping track of appointments, and performing email consultations. The app is available free of charge for all general practitioners (GPs) and their patients in Denmark [10,11]. The mobile app provided increased opportunity for use, and the number of VCs during the lockdown (March 2020) rose abruptly to 23,500 per week from approximately 0 before the lockdown. This number decreased to approximately 6000 VCs per week after the lockdown [12]. The uptake of VCs in general practice is dependent on a range of factors, and context-adequate guidance on how to implement and conduct a VC is one of many [3]. Developing guide materials may support those who plan for continued use and those who wish to embark on using this new consultation form beyond the COVID-19 period [12,13].

Objective

Research has demonstrated that a rapid cycle participatory design—involving stakeholders and end users in the development process—should be applied to ensure an efficient and agile approach. Using a participatory design provides a

unique opportunity for developing a guide based on stakeholders' and end users' knowledge and experiences [2,13,14]. This design approach may result in a quickly developed guide that is conducted in a realistic setup and that takes the opportunities and constraints that already exist within this specific context into consideration [15-17]. Moreover, research focused on developing tools and guides in a realistic setup adds further knowledge of methodological considerations needed to guide the development of future health care solutions [18].

The aim of this study is to rapidly develop a tool to assist general practice (ie, GPs and practice staff) in implementing VCs in daily practice. Therefore, we have developed a guide on how to implement VC in general practice in Denmark through a rapid cycle participatory design. In this study, we present the development process, methodological considerations, and evaluations made by general practice users.

Methods

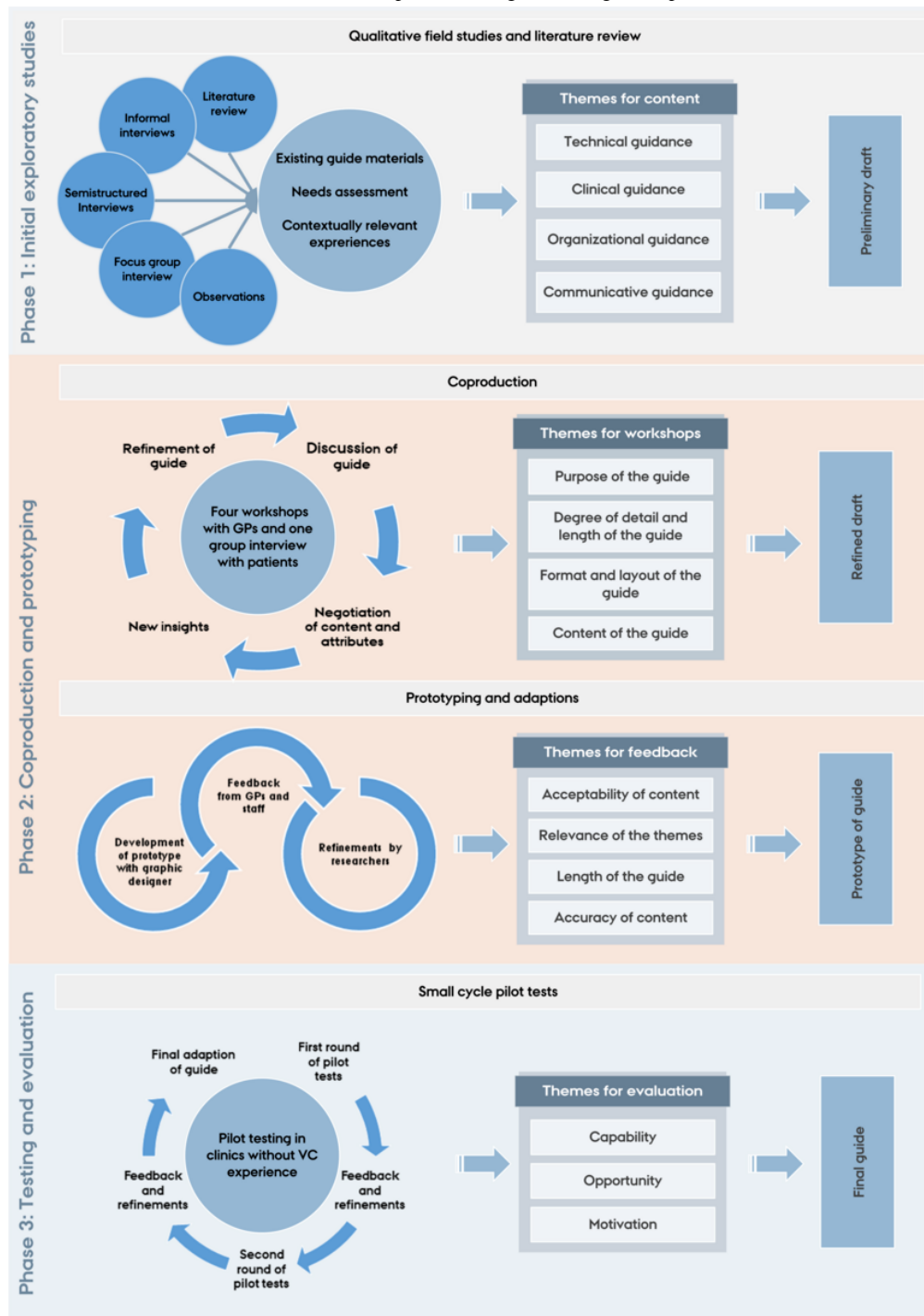
Setting

Danish health care is mainly funded by public taxes, with free-of-charge access to its services. General practice in Denmark is privately owned by GPs and organized into small units, either as single-handed practices (1 GP) or group practices (2-10 GPs), and almost every clinic has staff members. General practice is mostly financed through the public health care reimbursement scheme, and their services are regulated by collective agreements between the Danish Regions and the Organization of General Practitioners in Denmark [19,20]. At the beginning of the COVID-19 pandemic, a financial agreement was made for general practice to enable remuneration for the use of VCs [21]. General practice receives 165.27 DKK (US \$25,67) per VC.

Study Design

The development of the guide for general practice involved elements from coproduction approaches [22,23] and rapid cycle research [24]. This entails an approach in which researchers and end users collaborate throughout the project, using an iterative framework based on rapid qualitative analysis and interdisciplinary collaboration with a continuous exchange between research and practice. The study was undertaken in 3 phases comprising initial exploratory studies, coproduction, and evaluation, as illustrated in Figure 1. At different stages of the research process, insights from earlier phases were incorporated [25].

Figure 1. Overview of the framework and activities of the development of the guide. GP: general practitioner; VC: video consultation.



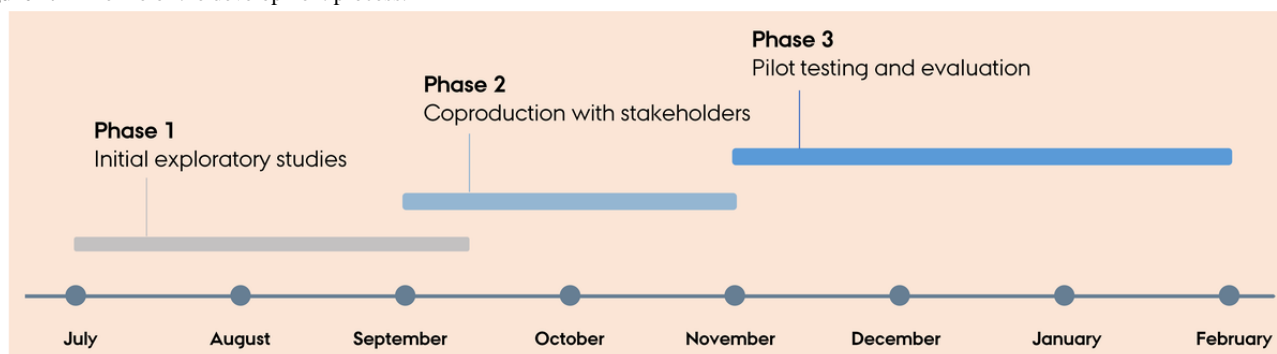
Analysis Team

The interdisciplinary analysis team conducting the ongoing analysis (AS, LDC, LLL, FGK, and NPC) comprised researchers with different backgrounds, including public health (AS and LLL), health science (NPC), medicine (FGK), and pharmaceutical sciences (LDC).

Phase 1: Initial Exploratory Studies

Phase 1 was conducted from July to September 2020 (Figure 2). We conducted a broad, explorative literature review,

searching both academic and gray literature [26] on existing telemedicine guides, including telephone consultations and VCs. A total of 17 guides were identified but included only if found suitable for the Danish general practice setting. For instance, guides were excluded if they were concerned with telecommunication exclusively between health care providers, focused on legal- or insurance-related aspects of telecommunication or specific technical programs, were too lengthy, or merely repeated points from other identified guides.

Figure 2. Timeline of the development process.

To identify the need for guidance, ensure relevance and acceptability of the guide, and reduce problems with implementation, we collected primary data through observations and informal and semistructured interviews with 7 full-time GPs, 3 practice staff, and 2 patients (performed by AS and LDC). The participants were identified from the authors' professional networks to foster a rapid first round of coproduction. Participants were not introduced to the research project before participation. Furthermore, a focus group interview was conducted with 7 data consultants, who were the technical or administrative staff charged with the role of providing technical and administrative support to general practice, for example, with VC implementation. The participants were purposefully sampled. This approach was used to select participants with broad experience in the field of VC. On the basis of the findings from the secondary and primary data, a first draft of the guide was developed.

Phase 2: Coproduction With Stakeholders

Overview

The coproduction phase entailed 4 workshops with GPs and 1 group interview [27] with patient representatives. Workshops

and group interviews were undertaken to develop the content of the draft guide developed in phase 1. The results from phase 2 were condensed into a prototype of the guide. The prototyping of the guide content was used to identify issues with acceptability and feasibility at an early point in development and allow them to be addressed before testing and evaluation [25].

Data Collection and Analysis

Phase 2 was conducted from September to November 2020 (Figure 2). The 4 workshops took place in 4 large cities in different regions in Denmark. The locations of the workshops in the large cities were chosen for convenience purposes but included GPs from across the country to ensure geographical variation. The participants were purposefully sampled based on their level of experience with VC. The number of participants ranged from 3 to 8 GPs in each workshop, which totaled 21 GPs (Table 1). All GPs had some or extensive experience with VCs. Workshops were used to discuss which key elements should be included in the guide to ease the implementation of VCs in general practice.

Table 1. Overview of participants (codesigners; N=54).

Type of participant	Phases	Number of participants	Gender (female), n (%)	Age interval, minimum-maximum
GPs ^a	1-3	32	16 (50)	36-71
Data consultants ^b	1	7	3 (43)	32-66
Practice staff	1-3	6	5 (83)	— ^c
Patients	1-2	4	1 (25)	39-58
National and regional stakeholders ^d	1	5	5 (100)	— ^c

^aGP: general practitioner.

^bTechnical or administrative staff with a supporting role to help general practice with technical issues. The data consultants are placed in different regions in Denmark.

^cNot possible.

^dLeaders or employees working in regional or national health institutions, that is, prehospital emergency medical services, regional quality development units, and general practice support units.

The workshops comprised individual and group activities commonly used in user experience design [28], such as reflection exercises, discussions, negotiations, and polls. All activities included graphic facilitation with active engagement, visualizations, or physical manifestations to provide a shared nexus for communication. The workshops were facilitated by

AS and LDC. The aim was to not only encourage participants to draw on their own VC experiences, creativity, and expertise but also to reflect on their own needs in the early implementation stage, as the end users of the guide would be inexperienced users. The facilitators led the GPs through 4 separate processes involving reflections on the barriers and facilitators related to

the use of VC in general practice, discussions on how to organize VCs in the existing general practice setting, and discussions on when to use VC instead of face-to-face consultation. Finally, the layout of the guide was discussed. Sufficient information power was obtained [29].

We recruited 2 patient representatives with significant knowledge on patient safety and patient–physician communication through the Danish Society for Patient Safety, and they participated in a group interview. They were asked to focus on the VC guide from their patient perspectives, and the themes and results of the GP workshops were presented to them. This allowed the patient representatives to provide inputs on how GPs and practice staff should implement VCs; for example, advice and knowledge regarding potential pitfalls in terms of ensuring patient safety and satisfaction with VC.

A total of 2 researchers (AS and LDC) from the analysis team and a research assistant reviewed the video recordings from the workshops, the participants' notes, and the researchers' notes from the workshops and group interviews. Subsequently, in collaboration, the researchers made iterations based on the open and axial coding of the participants' comments on the content in the guide, and new aspects were incorporated [30,31]. Data from one workshop were used to develop a preliminary codebook of predefined codes concerning the elements for the guide materials. On the basis of this codebook, each member of the analysis team analyzed parts of the workshop data and group interview data. Afterward, the analysis team discussed the identified themes and positions across the initial analyses to obtain consensus on the content of the guide, which was to be included in the new prototype of the guide [30,31]. The analysis team used NVivo (version 13; QSR International) software to manage and code all the data.

We invited former workshop participants, GPs, and practice staff from phase 1 (Table 1) to review the newly developed prototype of the guide in the form of a webpage. The review process was based on a rapid cycle improvement concept and included several small circles of reviewing of both the layout and components of the guide [32]. Participants were asked to assess 5 features: the acceptability of the guide materials in daily clinical practice, the relevance of the selected themes, the length of each section (considering the relevance of the particular theme), the accuracy of the information, and the format (whether the materials were presented in a convenient format with simple, clear, and easy-to-understand messages). A total of 2 researchers (AS and LDC) from the analysis team evaluated the statements and incorporated the adjustments that were widely agreed upon or, in other ways, deemed generic.

Phase 3: Pilot Testing and Evaluation

Phase 3 was conducted from November 2020 to February 2021 (Figure 2). After iterations of development and refinement, the prototype was tested by GPs and practice staff who had not been involved in the development phases and had little or no experience with VC. These participants were selected to represent the end users of the guide. A total of 7 semistructured interviews were undertaken by NPC. In total, 5 clinics from 4 different Danish regions pilot-tested the guide. The participants were purposefully sampled and had access to the materials for

approximately 2 months. During this period, they were asked to use the materials in daily practice and follow the proposed steps. Subsequently, 4 GPs and 3 practice staff members from these clinics participated in the evaluation interviews (Table 1). The interviews were conducted via telephone, audio recorded, and based on a semistructured interview guide that was inspired by the theory of capability-opportunity-motivation and behavior (COM-B) model, which proposes that people need capability (C), opportunity (O), and motivation (M) to perform a behavior (B) [33]. The COM-B model guided the structure of the interview guide into questions relating to capability (how and to what extent does the guide help GPs and practice staff obtain the required knowledge and skills in relation to VC implementation), opportunity (how and to what extent does the guide help to obtain and facilitate the use of time resources, equipment, and support properly), and motivation (how and to what extent does the guide help motivate VC implementation reflected through prioritization and emotional drive). Furthermore, the COM-B model was used as a theoretical framework for analyzing the data for the final revision of the guide materials [33]. Meaningful units related to the components of the COM-B model were identified and thematically coded. The analysis was also based on the rapid analysis approach [34], and iterative refinements were made during pilot testing [16]. Thematic coding was performed by AS, LDC, and NPC, and emerging themes were discussed with the other authors. This iterative process led to an agreement on the final version of the guide, which was adjusted upon completion of the pilot testing and evaluation phases.

Ethics

No approval from an ethical committee was needed for research involving observations, interviews, and workshops according to the national research guidelines in Denmark [35]. This study was registered in the Record of Processing Activities at the Research Unit for General Practice, Aarhus, in accordance with the provisions of the General Data Protection Regulation [36]. All participants from the workshops, focus group interviews, and face-to-face interviews gave written informed consent. Participants gave oral informed consent before participating in observations, informal interviews, or web-based interviews. The study complies with the Helsinki Declaration [37], and data storage and access comply with the General Data Protection Regulation.

Results

The results from the development of the guide have been described in the following subsections and illustrated in Figure 1.

Phase 1: Initial Exploratory Studies

We identified 17 documents containing guiding materials for VC from Denmark (11/17, 65%), the United States (1/17, 6%), Australia (1/17, 6%), and England (4/17, 24%). We included 3 guides [38-40] that were used as secondary sources for developing the content for the new guide in this study. This condensation was based on the considerations regarding relevance, acceptability, and potential to construct a meaningful guide. The primary data showed that GPs with minimal or no

experience were hesitant to use VC in practice, needed knowledge and experience with VC, experienced clinical, organizational, and technical barriers, were confused regarding where to begin, and requested guidance for implementing VCs. The data consultants expressed similar needs and barriers based on their knowledge of aiding GPs to start with VCs. In addition, data from VC-experienced GPs and staff helped identify the potential steps for the implementation of VCs related to workflow, organization, communication, and division of tasks. On the basis of these qualitative explorative studies and the existing literature, we identified 4 fundamental needs for guidance to successfully implement and promote VC in general practice: technical guidance, organizational guidance, communicative guidance, and clinical guidance. To meet these needs, we constructed 3 components inspired by the 3 international guides identified in the literature: a gross list of potential content themes, an initial prototype of the guide, and a *toolbox* comprising communication phrases and a list of medical conditions deemed relevant for VCs. All 3 components were to be considered and discussed in the following workshops (phase 2).

Phase 2: Coproduction and Prototyping

On the basis of a qualitative team coding analysis across the 4 GP workshops and the group interview with patient representatives, 7 main themes were constructed: *format of the guide, implementation of VC, organization, communication, competencies, technical issues, and medical conditions*. These main themes represent important aspects to consider when developing useful, relevant, and easy guiding materials.

Format of the Guide

The layout of the guide was a highly discussed topic during the workshops and was adjusted accordingly. GPs preferred short and concise guides without heavy text. They preferred the guide to not exceed 1 page. Some GPs suggested that the guide be presented as bullet points:

It is like when you buy a device at home, then it comes with a gigantic instruction manual that nobody reads. And then there are the five bullets in the "How to" instructions that everybody reads, and that is how it should be since both opportunities are required. But for those who are just about to start, there should be no more than 10 bullets. [GP 11]

Some GPs mentioned that they preferred colorful guides, and others emphasized that guides in the form of a postcard (hard copy 1-pager) were appealing. In addition, although a few GPs cherished the option of a hard copy edition, GPs generally agreed that the guide should be available on the web. The GPs had many suggestions for content, which contradicted the idea of a short format and argued for dividing the guide into a short and a long version. Some suggested a web-based guide with more detailed information and materials for the implementation of VC, whereas the hard copy should be in the form of a checklist.

Implementation

According to GPs, a VC guide should address both the one-off challenges expected to occur during implementation, such as

the technical aspects of setting up VCs, and the recurring challenges related to the development of routines, workflow, and division of labor. Moreover, GPs discussed how these barriers could be met during implementation. Some GPs experienced barriers related to the identification of relevant guidance, whereas others highlighted the significance of motivating practice staff for its use:

It is cooler when they [practice staff] take part in it, and are dedicated and can see the point of it. [GP 8]

Organization

The organizational aspects included structuring, planning, and coordination of activities related to VC. A substantial difference in the organizational structure was identified between the general practices, challenging the development of advice regarding a general, recommended organization of VC. GPs generally supported the idea of an introductory staff meeting before VC implementation, where the collaboration and workflow related to the implementation of VC could be discussed. Similarly, GPs endorsed the idea of a scheduled evaluation meeting in clinics, as a continuous exchange of experiences was not self-evident. Themes for both introductory and evaluation meetings were proposed and discussed:

I also believe it is important when implementing new things that you follow up on it. That is always the case, not only in video consultations, but I believe that it's very, very important, especially when we find something difficult, to evaluate and solve the problems that may otherwise make you decide not to use it. [GP 9]

Moreover, GPs raised concerns regarding the organization of new workflows for referring patients to VCs, as the practice staff was not confident regarding the new visitation possibilities. Thus, an example of how the visitation could be organized was incorporated into the guide.

Communication

The communicative theme comprised both 1-way (typically written) and 2-way (typically oral) communication among GPs, staff, and patients, for example, on the clinic's website and during consultation. Several GPs expressed that patient communication before and during the VC was challenging and requested guidance on this topic to ensure a safe environment for both the patient and the GP. According to both patients and GPs, patients also needed guidance before the consultation on, for example, booking, technical issues, preparation, and what to expect. In particular, the role of practice staff in communication with patients was mentioned as an important topic for the successful management of VCs:

When COVID-19 was at its worst, we had a lot of them [video consultations] every day. This number has gone down now, and I think that it is because the receptionist doesn't really remember it, and the patients don't really know that it's an option. [GP 10]

Competencies

It became apparent that practice staff has a key role in relation to the organizational aspects and the booking of VCs. Thus, the

continuous education of practice staff regarding their role and potential tasks in VC seemed advantageous. Although the GPs generally felt capable of performing a VC, several GPs benefited greatly from sharing experiences regarding the use of VC with other participants during the workshop. However, a *trial and error* approach was acknowledged as a common approach, and 1 GP mentioned that GPs often tried to solve the problems themselves, although guides were available. Moreover, a general opinion on the subject of *when and for whom VC is relevant* was that the best way to learn was through practical experience:

I need to throw myself into things and get some experience with a lot of things, where the mistakes may happen before I even realize what it concerns. If I do that [get knowledge about video consultation] in advance and don't get to use it until a month later, then I will have forgotten what I prepared anyway.
[GP 19]

Thus, GPs generally thought that enhancing the competencies for VC among GPs and the practice staff required more than a guide, as their competencies were determined by the amount of practice and time spent. However, nearly all the participating GPs recommended each practice to designate a VC *super user* among the employees in their clinic to aid colleagues in solving the problems encountered during the implementation and use of VCs.

Technical Issues

Most GPs expressed that they had experienced frustrating technical challenges when using the *My Doctor* app, such as problems regarding the booking of a VC, the web-based waiting room, the interaction between the medical software systems and the *My Doctor* solution, and the possibility for relatives to participate in a VC. Some said that technical challenges often constituted the main barrier to implementation, as most GPs did not consider technical skills a core task and thus easily lost their patience and turned toward familiar solutions, which would typically be face-to-face consultations:

If a general practice experiences problems with video consultation, they will quickly give up and never really get started. [GP 1]

Therefore, GPs suggested that a guide for implementation should address how to overcome the most common technical challenges and offer step-by-step advice on how to get through the technical challenges. Moreover, the guide should clearly state where to seek more information or help in the case of technical difficulties.

Medical Conditions

Medical conditions deemed relevant for VC and listed as part of the initial version of the guide were debated in the workshops. The relevance of VCs in general practice is dependent on a range of factors, which was also pointed out by the GPs in this study. Some GPs requested a list of conditions that were relevant for visitation to VC (for inspiration), whereas others expressed that a list of conditions suitable for VC would be too narrow and dichotomous. The GPs were concerned about compromising their professionalism by adhering to a list when planning for a VC, as it would never be comprehensive enough to encompass

all potential situations. Therefore, they suggested that visitation should always be an individual assessment. However, a significant barrier to the initiation of VCs was the identification of medical conditions suitable for a VC. A compromise agreed upon by the GPs was to recommend practices to start the VC implementation process by inviting well-known patients with relatively simple inquiries (eg, a status consultation with a familiar patient with well-regulated diabetes) to enable GPs to become gradually more comfortable with the use of VC without compromising patient safety. A GP explained it as follows:

We took the really easy patients and the really easy diagnoses, and then slowly added more and more.
[GP 8]

Patients' Perspectives

A total of 4 main points were extracted from the group interviews with the 2 patient representatives. First, they argued that the GP should not expect patients to book a VC as patients are not able to assess whether a VC would be suitable in their case and would, therefore, typically resort to the familiar form of consultation. Second, adequate patient communication is important to minimize misunderstandings and ensure patient safety. Third, the GP should inform the patient that a physical consultation would always be an option if necessary. Fourth, the GP should explain in detail to the patient what will happen during a VC. These results were incorporated into the guide.

Prototyping

On the basis of the workshops and group interviews, a prototype was developed in an iterative process with stakeholders. The process resulted in both linguistic and format corrections, which were decided by the analysis team. Elements that conflicted with daily clinical practice and technical details related to the setup for the *My Doctor* app were modified. The guide included sections on *how to get started* and *daily activities related to VC*. These were divided into 7 main steps: determining the organization of VCs, testing the technical setup of VCs, deciding how to use VCs, preparing patients to use VCs, performing a VC, learning from other GPs' experiences with VCs, and setting up a standard way of booking a VC. An additional section targeting the practice team was included in the guide; this section focused on how to evaluate the suitability of a patient's health problems for a VC and the administration of patient communications, appointments, and bookings. Several linguistic edits were made to reduce technical ambiguities. Finally, the guide was split into 2 documents: a full-length guide (webpage and downloadable file) and a checklist (downloadable file).

Phase 3: Testing and Evaluation

Interviews with GPs and practice staff conducted upon pilot testing showed that the clinics had used the materials differently. Some had followed the steps thoroughly and succeeded in performing several VCs with their patients, whereas others had implemented selected parts and conducted a few VCs. The following subsections will present how the GPs and staff experienced the guide's ability to equip them with the necessary capability, disclose pathways to acquire the opportunity, and thereby foster the motivation that drives the implementation process.

Capability—Strengthening Skills and Knowledge

After being introduced to the guide materials, both GPs and practice staff generally agreed that the guide had improved their knowledge regarding how to implement VC and thereby their capability of performing a VC:

It [the guide] was like following an easy instruction manual. Well, it was clearly written, it was divided into nice sections, which also made it easy if you were in doubt about something because then you could just go to that particular section. [GP 22]

However, GPs did not always use the materials as intended. It was common for the GPs to only read the checklist and instead rely on the practice staff to familiarize themselves with the combined materials (the webpage, guide, and checklist) because of time constraints. Several GPs pointed out that, by relying entirely on the checklist, they did not acquire the knowledge required to avoid frustrating challenges during the implementation process. During the interview session, some GPs recognized that they might have acquired the necessary knowledge base to avoid these frustrations if they had read the materials more thoroughly before performing the VCs. In cases where GPs found the guide materials insufficient, it appeared that the GPs were unaware of the relationship between the webpage, the guide, and the checklist. The limited knowledge of the relationship between the webpage, the guide, and the checklist could be a barrier to using the materials. In these situations, the overlapping materials (checklist and webpage) were experienced as redundant instead of helpful. However, both GPs who used the materials as intended and those who did not indicated that the guide had improved their overall capability of implementing and performing VCs in daily practice.

Accordingly, the practice staff who had used the entire materials expressed that the materials were sufficient and had conveyed the appropriate amount of knowledge for implementing VC. Furthermore, the guide was perceived to strengthen their skills in performing a VC, encouraging them to experience technical solutions from the patients' point of view. A practice staff explained the following:

And then it is a really good idea this thing about trying it yourself so that you can see it from the patients' point of view. This is extremely important. [Practice staff 1]

Although the guide helped most of the practice staff and GPs gain knowledge regarding how to invite patients for VC, it was still found difficult to obtain the skills in some practices, because patients often declined to participate in a VC:

Unfortunately, there are not many patients who want to have a VC. [Practice staff 1]

Opportunity—Facilitating Organization and Workflow

GPs and practice staff experienced that the guide could serve as help when organizing a VC and that it served as a tool to achieve the necessary structure in the clinic to be able to offer VCs, which contributed to increased implementation chances. The materials appealed to the readers at different experiential levels. For some clinics, it was the technical setup (eg, internet

speed and equipment) that was most vital for facilitating sufficient opportunities to implement VC, whereas others perceived the guidance to facilitate the organizational framework (eg, how to set up patient bookings in the booking system) as the most important part:

Well, the technical part of making it work and testing it was easy enough. It was first when we came to "how do we introduce this [VC]" that it became difficult. At that point it [the guide] turned out useful. [GP 23]

However, by nature, guides cannot facilitate the opportunity to ease the time pressure experienced in general practice. This challenge was addressed by many GPs, who reported a lack of time to become familiar with VC:

The guide is easy, but [creating] the routine is difficult. You have a busy schedule. [GP 23]

In addition, GPs were not economically compensated for the time they used to read the guide materials, which, according to some of the GPs, made it difficult to justify that reading the guide materials should be prioritized over other important tasks in daily practice. Consequently, GPs preferred starting quickly and spending little time reading before starting:

The guide was too long, and as a doctor, you just want to get started, right? And [you] don't have the time to read through the whole guide. [GP 24]

However, some clinics dealt with time constraints by letting one of the practice staff explore the materials more thoroughly. The practice staff member explained that she had used the guide to improve their clinic's home page, which increased the opportunity for VCs by increasing their accessibility to patients:

I needed some good advice on how to present the VCs at our homepage, and the guide was a big help. The GP used the checklist, and I used the long guide with the examples of phrases. [Practice staff 1]

In contrast, 1 GP mentioned that having read the guide thoroughly could have saved her time, as she would have avoided spending time on the consultations unfit for VC. She elaborated the following:

It is great that somebody has spent a lot of effort thinking about "Who is it [that is suited for video consultation]" and "think about whom you want to book," right?...I am a little like "no, let's just get started," and then some receptionists will deal with it [booking], and when they, for the fourth time, have booked a video consultation for somebody with a shoulder injury, then it's a little annoying because it's somewhat difficult to perform a physical examination of the shoulder through video. [GP 25]

Facilitating the opportunity for VC implementation was challenged by the fact that most GPs struggled with organizational issues (eg, delegating the task of opening the web-based waiting room in the *My Doctor* app). Although many GPs and practice staff explained that the guide had supported the establishment of VCs, an exact *recipe* for using the web-based waiting room was not offered in the guide. Instead,

it was put forward as an issue that the general practices should agree upon internally.

Motivation—Increasing the Motivation

Overall, the guide materials were perceived as acceptable, feasible, and relevant for the implementation of VC in general practice among both GPs and practice staff. The guide was described as useful, intuitive, and comprehensible. A common view was that it was easy to use and adopt in daily practice, which was described as a motivation to start:

Thank God, it [the guide] was easy. The guide made it much easier to get started with the VCs. [GP 22]

Moreover, the guide increased the motivation to implement VCs because of its simplicity (layout and content). In particular, the practice staff found it motivational to use the materials on the webpage, as it was found to be directly transferable and applicable in daily practice. One GP mentioned that the division of the guide into clear sections made it easily approachable. However, another GP argued that ideally, the guide materials should have been divided into sections that are aimed separately at GPs and practice staff.

Although the guide materials encouraged the introduction of team meetings in the implementation process, GPs described how they first and foremost felt that they had the responsibility. Some GPs felt alone during the process of implementing VCs. Combined with time limitations within the organization, this meant that the GPs easily lost their motivation, which made it difficult to change their daily routines:

I wish it could be possible to share more knowledge with colleagues about VCs—I feel alone with the responsibility to get the VCs on track. [GP 25]

Both GPs and staff expressed that being interviewed about the use of the guide and the checklist made them more motivated to start using VCs and use the guide more systematically for a good start. Taking the time to reflect on the guide materials acted as a motivational factor for the respondents. One GP explained that she would share her guide experiences with a colleague after the interview:

I will bring the checklist and the guide and all sorts of things to my colleague tomorrow so that she can try to get started too. [GP 23]

GPs and the practice staff agreed that they would recommend the guide materials to other colleagues, and they thought that the guide had addressed the most important issues regarding the implementation of VC. Thus, in addressing capability and opportunity, the guide increased the motivation for implementing VC.

Final Adaptions

On the basis of the feedback in phase 3, minor adjustments were made to wording and format. The contents of the guide were then finalized. The final 7 main themes identified were as follows: determining the organization of VC, testing the technical setup of VC, deciding how to use VC, preparing patients to use VC, performing a VC, arranging bookings for VCs, and considering other useful advice [41].

Some GPs found that the guide comprised sections irrelevant to their specific situation, which ultimately made it too lengthy. This experience of relevance varied between GPs and across sections, making it difficult to shorten the guide. Therefore, we inserted a table of contents and customized the layout with clear headings to enable future users to easily decide upon the sections they found pertinent.

The pilot test indicated that GPs and practice staff were capable of adopting the VC guide to implement VC in general practice. In addition, a short introductory video was added to the webpage to underpin the relationship between the guide materials and support the implementation process.

Discussion

Principal Findings

A rapid cycle coproduction approach was used to explore the needs of general practice concerning the implementation of VCs. This approach was taken to develop useful, relevant, and easy guiding materials. The 3-phased framework used in this study provides a pragmatic example of coproducing and prototyping tools that fit into the everyday clinical workflow and meet the needs of general practice. The use of a small cycle iterative approach enabled us to address implementation issues at the design stage and modify the tool to overcome common barriers [25]. Our framework offers insight into how collaboration and coproduction with stakeholders can be incorporated into these different stages of intervention development. In codesign approaches, end users are involved throughout the development processes and work together with the research team during all phases [17,42]. In this study, GPs, practice staff, patients, and other stakeholders contributed to the development by generating content specific to the guide.

During data analysis and guide refinement, a rapid assessment procedure was used. This approach entails an in-depth understanding of the important elements for the guide without transcribing verbatim. This is a useful approach for creating comprehensive information in short timelines [43]. Our approach was similar to other rapid methods, such as using a priori structured codebook [44], coding audio only [45], and allocating researchers to code for specific themes [46]. Such rapid analysis approaches have been shown to produce valid findings, compared with traditional in-depth, line-by-line transcript analysis [34]. Thus, when conducting research with time constraints, this approach could be considered a supplement to the qualitative researcher toolkit [44].

The COM-B behavioral-inspired analysis [33] gave insights into how the guide was adapted and used in a clinical setting. GPs and practice staff felt equipped to implement VC after reading the guide. However, to implement VC in daily practice, the GPs needed to spend some time organizing VC in general practice. However, the guide motivated GPs and practice staff to implement VC. However, this implementation will not succeed solely by distributing the guide. Following the diffusion of innovation theory [47] and as widely recognized in the empirical literature, some clinicians adopt technological innovations readily, whereas others need motivation and support

[13]. The guide in itself may not facilitate GPs gaining experience and confidence in VCs. Instead, GPs who are, by nature, not early adopters would benefit from a multifaceted intervention to aid the implementation of VC (eg, through peer training and the possibility of debating with colleagues having more extensive VC experience). Jensen et al [48] concluded that GPs need different approaches to implementation support and that the combination of support types depends on their needs and willingness to invest resources in future interventions. Moreover, a study from Australia [49] concluded that establishing VC as routine practice would need to be endorsed by patients, GPs, and founding organizations. Thus, proper complementary supportive interventions would be required to obtain the synergic effects of a multifaceted approach and fulfill the potential of the guide materials.

Resistance to change is common, and concerns relating to the introduction of alternative or supplementary methods of consultation can be expressed as concerns regarding patient safety. GPs and practice staff have responded to alternative forms of consultation with a mix of enthusiasm for innovation and resistance to change [7]. However, a guide developed with stakeholder engagement through coproducing methods could infuse less resistance, as peer professionals are listed as senders of the messages. Moreover, the guide addresses difficulties and concerns that have been empirically defined. However, the implementation is likely to be a difficult and resource-intensive task that would require both national and local strategic leadership [50].

The guiding materials were developed during the COVID-19 pandemic. These contextual factors may have influenced the development process, as many factors changed rapidly during the process [17]. This could affect the topicality of the guide during the postpandemic period. However, by having paid attention to the common usability of the guide across time (eg, by informing stakeholders to consider the VC guide from a general and decontextualized perspective), we consider the guide to be relevant beyond the COVID-19 context. Nonetheless, the guide is both applicable and particularly relevant in a COVID-19 context, as VCs may reduce the risk of COVID-19 contagion because of the reduced physical contact.

The use of VC during the COVID-19 pandemic may yield useful knowledge regarding which patients and diseases would be relevant for VC and whether the use of VCs may be extended to a wide variety of patients and clinical situations. Moreover, there is a long-term perspective on the use of VC and thus a strong need for a guide on its implementation. The population of older adults is increasing, and the aging population has an increasing number of health problems. This may further increase the pressure on primary health care, which is often the first point of contact. In Denmark and similar countries, access to GPs is a rising concern, especially as their unavailability may push some people to seek help elsewhere or not at all [51,52]. If quicker and easier access to care means that more people in need of care would contact general practice, VC could be adopted as a patient entitlement under universal health coverage [53]. However, offering VCs may also increase the workload in general practice unless the video-based encounter is of shorter duration than (or of similar length to) an ordinary face-to-face

consultation and requires no subsequent consultation [53]. The provision of VC is likely to depend on the remuneration for this type of service, and the remuneration of GPs for services provided is regulated by collective agreements between the Danish Regions and the Danish Organization of General Practitioners in Denmark [19].

Strengths and Limitations

A strength of this study is that all steps in the development process were described in detail. Using a coproduction approach with the triangulation of multiple data sources was an additional strength. Coproduction is known to improve the adaptation of the guide to a given context [25]. Using a coproduction approach allowed us to involve few but specifically selected participants for each phase. Concurrently, the rapid cycle approach ensured an agile process with adequate and suitable inputs.

Analysis of the qualitative data with a team of coders proved highly successful in terms of reaching as many different perspectives on the available data as possible. The inclusion of additional coders and their interpretations served to enhance the credibility of the framework that emerged.

In 2 out of 4 workshops, it was only possible to recruit 3 participating GPs in each workshop compared with the 8 and 7 participating GPs in the 2 other workshops. Furthermore, although the 2 workshops contained only 3 GPs, we assessed that they contributed with just as many diverse perspectives and nuances regarding the content of the guide as those provided in the workshops with more participants. To some degree, homogeneity is recommended for participatory design [14]. The heterogeneity in our workshops showed that GPs had different needs and preferences. Generally, the coproduction processes contributed to identifying many perspectives on and experiences with VC to improve the guide's content.

General practice in Denmark is organized into different types of clinics. VC will be implemented and used differently in the clinics, and it can be difficult to adapt VC to all the clinics. However, in all phases of this study, the recruited GPs were initially selected to represent a variety of GPs with and without experience with VC [17]. All GPs and other stakeholders were purposely sampled across the nation. Another strength of the developed materials is that the implementation guide and the website are freely available in a format that is directly applicable to general practice.

A limitation was that we included only 2 patients in the analysis. Consequently, we must conclude that the patient perspective is only sparsely presented in the guide. However, the included patients were formally appointed patient representatives who were experienced in speaking on behalf of a broad patient population. Moreover, the documentation was challenged in the rapid analysis, which limited the controllability of the analysis, as the usual coding process was reduced to an oral discussion [54]. We addressed this by recording and revisiting the group discussion to enable tracking of the iterative analysis. The rapidity of our analysis implied that we did not progress into higher levels of abstraction and interpretation but rather focused on the potential improvements to the guide.

Conclusions

Coproduction, involving prototyping, small iterations, and rapid data analysis, is a suitable approach when contextually rich, *hands-on* guide materials are urgently needed. The new guide for VCs was developed through rapid cycles, team-based work,

and acknowledged research methodology. It comprised the following themes: organizing a VC in general practice, testing the technical setup of VC, deciding how to use VC, preparing patients to use VCs, performing a VC, arranging bookings for a VC, and proposing useful advice.

Acknowledgments

The authors thank the general practitioners, practice staff, patients, and data consultants for their participation in the development of the guide. Furthermore, the authors would also like to acknowledge the Danish Society for Patient Safety for knowledge exchange regarding the rapid cycle improvement concept and the European Society for Quality and Safety in Family Practice for the graphic setup of the guide. In addition, the authors thank Marlene CR Møller for the design and development of workshop protocols, Sine Hammer Hansen for project assistance during phases 1 and 2, Lone Niedziella for proofreading, and Ulrik Bak Kirk for project management. Finally, the authors would like to thank the principal investigator Per Kallestrup and the steering group comprising Jens Søndergaard, Susanne Reventlow, and Janus L Thomsen for guidance, counseling, and support.

This work was supported by Innovation Fund Denmark. The funding source had no role in the design, analysis, and interpretation of the data or in the writing of the paper.

Authors' Contributions

AS and LDC contributed to the study design, data collection, data analysis, and manuscript writing. NPC and LLL contributed to data collection, data analysis, and manuscript writing. FGK contributed to data analysis and manuscript writing. ADG, CHM, and EAH contributed with support to the data analysis and manuscript writing. All the authors have read, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization Director - General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. 2020. URL: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> [accessed 2021-10-30]
2. Heltberg J. Videokonsultationer skal mindske smitterisikoen hos de praktiserende læger [Video consultations are to reduce the risk of contagion among GPs]. The Danish Organization of General Practitioners. 2020. URL: <https://www.laeger.dk/nyhed/videokonsultationer-skal-mindske-smitterisikoen-hos-de-praktiserende-laeger> [accessed 2021-10-30]
3. 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):bjgpopen20X1010. [doi: [10.3399/bjgpopen20X101020](https://doi.org/10.3399/bjgpopen20X101020)] [Medline: [32184212](https://pubmed.ncbi.nlm.nih.gov/32184212/)]
4. Atherton H, Brant H, Ziebland S, Bikker A, Campbell J, Gibson A, et al. Alternatives to the face-to-face consultation in general practice: focused ethnographic case study. *Br J Gen Pract* 2018 Apr;68(669):293-300 [FREE Full text] [doi: [10.3399/bjgp18X694853](https://doi.org/10.3399/bjgp18X694853)] [Medline: [29378697](https://pubmed.ncbi.nlm.nih.gov/29378697/)]
5. Donaghy E, Atherton H, Hammersley V, McNeilly H, Bikker A, Robbins L, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. *Br J Gen Pract* 2019 Sep;69(686):586-594 [FREE Full text] [doi: [10.3399/bjgp19X704141](https://doi.org/10.3399/bjgp19X704141)] [Medline: [31160368](https://pubmed.ncbi.nlm.nih.gov/31160368/)]
6. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. *Fam Med Community Health* 2020 Aug;8(3):e000530 [FREE Full text] [doi: [10.1136/fmch-2020-000530](https://doi.org/10.1136/fmch-2020-000530)] [Medline: [32816942](https://pubmed.ncbi.nlm.nih.gov/32816942/)]
7. Atherton H, Ziebland S. What do we need to consider when planning, implementing and researching the use of alternatives to face-to-face consultations in primary healthcare? *Digit Health* 2016;2:2055207616675559 [FREE Full text] [doi: [10.1177/2055207616675559](https://doi.org/10.1177/2055207616675559)] [Medline: [29942570](https://pubmed.ncbi.nlm.nih.gov/29942570/)]
8. Tæt på borgerens hverdag [Close to the citizen's everyday life]. The Danish Regions. URL: <https://www.regioner.dk/sundhed/digitalt-sundhed-for-dig/taet-paa-borgerens-hverdag> [accessed 2021-10-30]
9. Melbye T. Praktiserende læger i Region Sjælland fravælger video [General practitioners in Region Zealand opt out of video]. *Dagens Medicin*. 2020. URL: <https://dagensmedicin.dk/praktiserende-laeger-i-region-sjaelland-fravaelger-video/> [accessed 2021-10-30]
10. The My Doctor app has become the most popular in Denmark. *Trifork*. 2021. URL: <https://trifork.com/2020/04/doctor-app-has-become-the-most-popular-in-denmark/> [accessed 2021-10-30]

11. Video consultations are now live in the MyDoctor App (MinLæge) in Denmark. Trifork. 2021. URL: <https://trifork.com/2020/03/video-consultations-are-now-live-in-the-mydoctor-app/> [accessed 2021-10-30]
12. Almen praksis under COVID-19: Hurtig omstilling i almen praksis sikrede danskerne lægehjælp på video og telefon i starten af coronakrisen [[General practice during COVID-19: Rapid adjustment in general practice ensured provision of medical care through video and telephone at the beginning of the corona crisis]]. The Danish Organization of General Practitioners. 2020. URL: <https://www.laeger.dk/sites/default/files/coronaanalyseendeligtoktober.pdf> [accessed 2021-10-30]
13. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. *J Med Internet Res* 2018 Dec 17;20(4):e150 [FREE Full text] [doi: [10.2196/jmir.9897](https://doi.org/10.2196/jmir.9897)] [Medline: [29625956](https://pubmed.ncbi.nlm.nih.gov/29625956/)]
14. Kushniruk A, Nøhr C. Participatory design, user involvement and health it evaluation. *Stud Health Technol Inform* 2016;222:139-151. [Medline: [27198099](https://pubmed.ncbi.nlm.nih.gov/27198099/)]
15. Shah SGS, Robinson I. User involvement in healthcare technology development and assessment: structured literature review. *Int J Health Care Qual Assur Inc Leadersh Health Serv* 2006;19(6-7):500-515. [doi: [10.1108/09526860610687619](https://doi.org/10.1108/09526860610687619)] [Medline: [17100220](https://pubmed.ncbi.nlm.nih.gov/17100220/)]
16. Etchells E, Ho M, Shojania KG. Value of small sample sizes in rapid-cycle quality improvement projects. *BMJ Qual Saf* 2016 Mar;25(3):202-206. [doi: [10.1136/bmjqs-2015-005094](https://doi.org/10.1136/bmjqs-2015-005094)] [Medline: [26717989](https://pubmed.ncbi.nlm.nih.gov/26717989/)]
17. 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/)]
18. Vandekerckhove P, de Mul M, Bramer WM, de Bont AA. Generative participatory design methodology to develop electronic health interventions: systematic literature review. *J Med Internet Res* 2020 Apr 27;22(4):e13780 [FREE Full text] [doi: [10.2196/13780](https://doi.org/10.2196/13780)] [Medline: [32338617](https://pubmed.ncbi.nlm.nih.gov/32338617/)]
19. Pedersen KM, Andersen JS, Søndergaard J. General practice and primary health care in Denmark. *J Am Board Fam Med* 2012 Mar;25 Suppl 1:34-38 [FREE Full text] [doi: [10.3122/jabfm.2012.02.110216](https://doi.org/10.3122/jabfm.2012.02.110216)] [Medline: [22403249](https://pubmed.ncbi.nlm.nih.gov/22403249/)]
20. Beskrivelse af almen praksis sektoren i Danmark [Description of the general practice sector in Denmark]. The Danish Health Data Authority, Copenhagen. 2020. URL: <https://sundhedsdatastyrelsen.dk/da/tal-og-analyser/analyser-og-rapporter/almen-praksis-og-kommuner/almen-praksis> [accessed 2021-10-30]
21. Aftaler i almen praksis under coronaepidemien [Agreement in general practice during the corona epidemic]. The Danish Organization of General Practitioners. URL: <https://www.laeger.dk/PLO/aftaler-i-almen-praksis-under-coronaepidemien#centrale> [accessed 2021-11-12]
22. Vennik FD, van de Bovenkamp HM, Putters K, Grit KJ. Co-production in healthcare: rhetoric and practice. *Int Rev Adm Sci* 2015 Jun 29;82(1):150-168. [doi: [10.1177/0020852315570553](https://doi.org/10.1177/0020852315570553)]
23. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. *BMJ Qual Saf* 2016 Jul;25(7):509-517 [FREE Full text] [doi: [10.1136/bmjqs-2015-004315](https://doi.org/10.1136/bmjqs-2015-004315)] [Medline: [26376674](https://pubmed.ncbi.nlm.nih.gov/26376674/)]
24. Schneeweiss S, Shrank WH, Ruhl M, Maclure M. Decision-making aligned with rapid-cycle evaluation in health care. *Int J Technol Assess Health Care* 2015 Jan;31(4):214-222. [doi: [10.1017/S0266462315000410](https://doi.org/10.1017/S0266462315000410)] [Medline: [26585440](https://pubmed.ncbi.nlm.nih.gov/26585440/)]
25. Hawkins J, Madden K, Fletcher A, Midgley L, Grant A, Cox G, et al. Development of a framework for the co-production and prototyping of public health interventions. *BMC Public Health* 2017 Sep 04;17(1):689 [FREE Full text] [doi: [10.1186/s12889-017-4695-8](https://doi.org/10.1186/s12889-017-4695-8)] [Medline: [28870192](https://pubmed.ncbi.nlm.nih.gov/28870192/)]
26. Adams J, Hillier-Brown FC, Moore HJ, Lake AA, Araujo-Soares V, White M, et al. Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. *Syst Rev* 2016 Sep 29;5(1):164 [FREE Full text] [doi: [10.1186/s13643-016-0337-y](https://doi.org/10.1186/s13643-016-0337-y)] [Medline: [27686611](https://pubmed.ncbi.nlm.nih.gov/27686611/)]
27. Frey JH, Fontana A. The group interview in social research. *J Soc Sci* 2019 Dec 09;28(2):175-187. [doi: [10.1016/0362-3319\(91\)90003-m](https://doi.org/10.1016/0362-3319(91)90003-m)]
28. Gibbons S. UX mapping methods compared: a cheat sheet. Nielsen Norman Group. 2017. URL: <https://www.nngroup.com/articles/ux-mapping-cheat-sheet/> [accessed 2021-10-30]
29. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res* 2016 Nov;26(13):1753-1760. [doi: [10.1177/1049732315617444](https://doi.org/10.1177/1049732315617444)] [Medline: [26613970](https://pubmed.ncbi.nlm.nih.gov/26613970/)]
30. Devotta K, Pedersen C. Coding qualitative data: working with a team of coders. CRICH Survey Research Unit Methodology Bits. 2015. URL: <https://tspace.library.utoronto.ca/handle/1807/70247> [accessed 2021-10-30]
31. Richards KA, Hemphill MA. A practical guide to collaborative qualitative data analysis. *J Teach Phys Educ* 2018 Apr;37(2):225-231. [doi: [10.1123/jtpe.2017-0084](https://doi.org/10.1123/jtpe.2017-0084)]
32. Norman CL, Nola TW, Moen R, Provost LP, Nolan KM, Langley GJ. *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*. Hoboken, New Jersey, United States: John Wiley & Sons, Inc; 2014.
33. 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/)]

34. Vindrola-Padros C, Vindrola-Padros B. Quick and dirty? A systematic review of the use of rapid ethnographies in healthcare organisation and delivery. *BMJ Qual Saf* 2018 Apr 20;27(4):321-330. [doi: [10.1136/bmjqs-2017-007226](https://doi.org/10.1136/bmjqs-2017-007226)] [Medline: [29263139](https://pubmed.ncbi.nlm.nih.gov/29263139/)]
35. Act on research ethics review of health research projects - Pub. L. No. 593. The Danish National Committee on Health Research Ethics. 2011. URL: <https://en.nvk.dk/rules-and-guidelines/act-on-research-ethics-review-of-health-research-projects> [accessed 2021-10-30]
36. Complete guide to GDPR compliance. Proton Technologies AG. 2021. URL: <https://gdpr.eu/> [accessed 2021-10-30]
37. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Med Assoc* 2013 Nov 27;310(20):2191-2194. [doi: [10.1001/jama.2013.281053](https://doi.org/10.1001/jama.2013.281053)] [Medline: [24141714](https://pubmed.ncbi.nlm.nih.gov/24141714/)]
38. Telephone and video consultations in general practice: flowcharts. Royal Australian College of General Practitioners. 2020. URL: <https://www.racgp.org.au/clinical-resources/covid-19-resources/telehealth/phone-and-video-consultations-in-general-practice> [accessed 2021-10-30]
39. A checklist for incorporation of video visits (telemedicine). American College of Physicians. 2020. URL: https://www.acponline.org/system/files/documents/practice-resources/health-information-technology/telehealth/video_visit_telemedicine_checklist_web.pdf [accessed 2021-10-30]
40. Greenhalgh T. Video consultations: a guide for practice. *BJGP Life*. 2020. URL: <https://bjgplife.com/video-consultations-guide-for-practice/> [accessed 2021-10-30]
41. Kom godt i gang med videokonsultation [Get started with video consultation]. The Research Units for General Practice in Denmark. 2021. URL: <https://remoteguide.dk/> [accessed 2021-10-30]
42. Morris RL, Ruddock A, Gallacher K, Rolfe C, Giles S, Campbell S. Developing a patient safety guide for primary care: a codesign approach involving patients, carers and clinicians. *Health Expect* 2021 Mar;24(1):42-52 [FREE Full text] [doi: [10.1111/hex.13143](https://doi.org/10.1111/hex.13143)] [Medline: [33142022](https://pubmed.ncbi.nlm.nih.gov/33142022/)]
43. Palinkas LA, Zatzick D. Rapid assessment procedure informed clinical ethnography (RAPICE) in pragmatic clinical trials of mental health services implementation: methods and applied case study. *Adm Policy Ment Health* 2019 Mar 28;46(2):255-270 [FREE Full text] [doi: [10.1007/s10488-018-0909-3](https://doi.org/10.1007/s10488-018-0909-3)] [Medline: [30488143](https://pubmed.ncbi.nlm.nih.gov/30488143/)]
44. Gale RC, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the Veterans Health Administration. *Implement Sci* 2019 Feb 01;14(1):11 [FREE Full text] [doi: [10.1186/s13012-019-0853-y](https://doi.org/10.1186/s13012-019-0853-y)] [Medline: [30709368](https://pubmed.ncbi.nlm.nih.gov/30709368/)]
45. Neal JW, Neal ZP, VanDyke E, Kornbluh M. Expediting the analysis of qualitative data in evaluation. *Am J Eval* 2014 May 27;36(1):118-132. [doi: [10.1177/1098214014536601](https://doi.org/10.1177/1098214014536601)]
46. Ash JS, Sittig DF, McMullen CK, Guappone K, Dykstra R, Carpenter J. A rapid assessment process for clinical informatics interventions. *AMIA Annu Symp Proc* 2008 Nov 06;2008:26-30 [FREE Full text] [Medline: [18999075](https://pubmed.ncbi.nlm.nih.gov/18999075/)]
47. Rogers EM. Diffusion of Innovations, 4th Edition. New York: Free Press; 2010.
48. Jensen ML, Bro F, Mygind A. Implementation of healthcare accreditation in Danish general practice: a questionnaire study exploring general practitioners' perspectives on external support. *Scand J Prim Health Care* 2021 Mar;39(1):85-91 [FREE Full text] [doi: [10.1080/02813432.2021.1882084](https://doi.org/10.1080/02813432.2021.1882084)] [Medline: [33646089](https://pubmed.ncbi.nlm.nih.gov/33646089/)]
49. Jiwa M, Meng X. Video consultation use by Australian general practitioners: video vignette study. *J Med Internet Res* 2013 Jun 19;15(6):e117 [FREE Full text] [doi: [10.2196/jmir.2638](https://doi.org/10.2196/jmir.2638)] [Medline: [23782753](https://pubmed.ncbi.nlm.nih.gov/23782753/)]
50. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for COVID-19. *Br Med J* 2020 Mar 12;368:m998. [doi: [10.1136/bmj.m998](https://doi.org/10.1136/bmj.m998)] [Medline: [32165352](https://pubmed.ncbi.nlm.nih.gov/32165352/)]
51. GP patient survey. National Health Service. 2019. URL: <http://www.gp-patient.co.uk/surveysandreports> [accessed 2021-10-30]
52. Lægedækning i hele Danmark [GP coverage in Denmark]. The Danish Health Authority. 2020 Mar 03. URL: <https://sum.dk/publikationer/2020/marts/anden-status-paa-laegedaekning-i-hele-danmark> [accessed 2021-10-30]
53. Salisbury C, Murphy M, Duncan P. The impact of digital-first consultations on workload in general practice: modeling study. *J Med Internet Res* 2020 Jun 16;22(6):e18203 [FREE Full text] [doi: [10.2196/18203](https://doi.org/10.2196/18203)] [Medline: [32543441](https://pubmed.ncbi.nlm.nih.gov/32543441/)]
54. Anfara VA, Brown KM, Mangione TL. Qualitative analysis on stage: making the research process more public. *Edu Res* 2016 Jul 01;31(7):28-38. [doi: [10.3102/0013189x031007028](https://doi.org/10.3102/0013189x031007028)]

Abbreviations

- COM-B:** capability-opportunity-motivation and behavior
 - GP:** general practitioner
 - VC:** video consultation
-

Edited by G Eysenbach; submitted 27.05.21; peer-reviewed by H Atherton, J Newbould, E Donaghy, J Hawkins; comments to author 21.07.21; revised version received 06.09.21; accepted 05.10.21; published 24.11.21.

Please cite as:

Sandbæk A, Christensen LD, Larsen LL, Primholdt Christensen N, Kofod FG, Guassora AD, Merrild CH, Assing Hvidt E
Guidance for Implementing Video Consultations in Danish General Practice: Rapid Cycle Coproduction Study

JMIR Form Res 2021;5(11):e27323

URL: <https://formative.jmir.org/2021/11/e27323>

doi: [10.2196/27323](https://doi.org/10.2196/27323)

PMID: [34821560](https://pubmed.ncbi.nlm.nih.gov/34821560/)

©Amanda Sandbæk, Line Due Christensen, Lotte Lykke Larsen, Nina Primholdt Christensen, Frida Greek Kofod, Ann Dorrit Guassora, Camilla Hoffmann Merrild, Elisabeth Assing Hvidt. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 24.11.2021. 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 a Wearable Biocueing App (Sense-IT) Among Forensic Psychiatric Outpatients With Aggressive Behavior: Design and Evaluation Study

Annemieke ter Harmsel^{1,2,3}, MSc, MA; Thimo van der Pol^{1,2,3}, PhD; Lise Swinkels^{1,2,3}, MSc; Anna E Goudriaan^{3,4}, ProfDr; Arne Popma², ProfDr; Matthijs L Noordzij⁵, PhD

¹Inforsa, Forensic Mental Health Care, Amsterdam, Netherlands

²Department of Child and Adolescent Psychiatry and Psychology, Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

³Department of Research and Quality of Care, Arkin, Amsterdam, Netherlands

⁴Department of Psychiatry, Amsterdam Public Health Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

⁵Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

Corresponding Author:

Annemieke ter Harmsel, MSc, MA

Inforsa, Forensic Mental Health Care

Vlaardingenvlaan 5

Amsterdam, 1059 GL

Netherlands

Phone: 31 653810776

Email: annemieke.ter.harmsel@inforsa.nl

Abstract

Background: The ability to regulate anger is often impaired in forensic psychiatric patients, frequently resulting in aggressive behavior. Although some treatment programs are partially successful in enhancing aggression regulation and reducing recidivism among specific subgroups, generalizable conclusions on the effectiveness of these interventions cannot be drawn to date. In forensic outpatient care, low treatment adherence and a predominant focus on cognitive control in most treatment programs may entail some of the factors impeding treatment. Technology-based interventions may address some of these treatment challenges.

Objective: The aim of this study is to explore whether a new technology-based biocueing intervention, the Sense-IT app, can be a valuable addition to aggression regulation treatment programs in forensic outpatient care. The Sense-IT app, which provides the user with real-time physiological feedback and behavioral support, is developed to strengthen emotional awareness and facilitate real-life practice. In this study, we aim to develop and evaluate an updated version of the Sense-IT app that is suitable for forensic outpatients with aggressive behavior.

Methods: First, we conducted a design study to assess the attitudes of forensic professionals and patients toward biocueing and to collect requirements for a biocueing app for this specific population. On the basis of this information, we developed an updated version of the Sense-IT app. In an evaluation study, 10 forensic outpatients used the app for 2 weeks. The app's acceptability, usability, and clinical outcomes (aggression, anger, and recognition of bodily signals related to anger) were measured before and after the intervention using both quantitative and qualitative measures.

Results: The design study revealed a cautiously positive attitude toward the use of biocueing as an addition to aggression regulation therapy. The evaluation study among forensic outpatients demonstrated moderate acceptability and adequate usability for the new version of the Sense-IT app. Exploratory analysis revealed a significant decrease in trait aggression postintervention; no significant changes were found in other anger-related clinical outcomes. To further increase acceptability and usability, a stable functioning app with self-adjustable settings, the use of smartwatches with a longer battery life, and the use of the patient's own smartphone devices were recommended.

Conclusions: This study, which is one of the first attempts to enroll and evaluate the real-life use of a biocueing intervention among forensic outpatients, emphasized the importance of involving both patients and therapists throughout the development and implementation process. In the future, experimental studies, including single-case experimental designs using ecological momentary assessment, should be performed to evaluate the effectiveness of the Sense-IT intervention on clinical outcomes. An

open attitude toward new technology, allowing exploration of the potential benefits of the Sense-IT app case-by-case, and training of therapists in using the app are expected to facilitate its integration in therapy.

(*JMIR Form Res* 2021;5(11):e29267) doi:[10.2196/29267](https://doi.org/10.2196/29267)

KEYWORDS

biocueing; biosensing; biofeedback; aggression; forensic psychiatry; wearable technology; mobile apps; usability; evaluation; mHealth

Introduction

Background

Aggression and violent behavior are associated with substantial problems, especially for the victims and the offenders. While victims of violence are at a high risk of developing psychological and behavioral problems such as depression, anxiety, posttraumatic stress, and alcohol abuse [1,2], offenders show increased rates of developing psychiatric disorders, such as psychotic, mood, and substance use disorders [3,4]. In addition, the offenders' imprisonment often leads to more unfavorable situations for them after release, such as job loss, housing problems, and a lack of social support. Consequently, aggressive behavior is a high burden for professionals working in the mental health care and the judicial system, resulting in inflated costs for both health care and the society in general [5,6]. The impact of aggressive behavior on both individual lives and society in general highlights the importance of early and effective treatment for forensic psychiatric patients with problematic aggressive behavior. Given the current trend of preferring outpatient interventions over residential treatment, the importance of real-life, out-of-session practice of behavioral alternatives [7], and the developments in the use of digital technology in psychological treatment [8], technology-based interventions are of interest to improve and support aggression regulation among forensic (outpatient) populations.

Over the years, several therapeutic interventions have been developed to reduce aggressive behavior and criminal recidivism among forensic in- and outpatients. Most offender treatment programs are based on the principles of aggression replacement training (ART) [9], in which behavioral, affective, and cognitive components are combined to improve anger and aggression regulation. In forensic psychiatry, these cognitive behavioral therapy (CBT) programs are broadly considered as promising rehabilitative treatments for antisocial behavior [10,11]. However, mixed findings have been reported regarding the effectiveness of these treatment programs. A meta-analysis of 14 CBT-informed anger management studies revealed an overall 28% risk reduction in violent recidivism after treatment, with a 56% reduction among those who completed the treatment. In half of the studies, significant differences in violent reoffense were reported, when compared with control conditions [12]. In a systematic review, including 16 ART studies, researchers stated that generalizable conclusions could not be drawn owing to, among other factors, differences in the severity of psychopathology, low methodological quality, and limited follow-up information [13].

The limited effectiveness of current treatment programs is also related to challenges specific to forensic populations. First,

because of severe psychopathology, a lack of problem awareness, and motivational difficulties, treatment engagement and adherence are often low. Among forensic populations, the rates of pretreatment drop out and treatment attrition are high. A large meta-analysis revealed an overall treatment attrition rate of 27.1% across all offender programs, with a rate of 37.8% for domestic violence offenders [14]. Research results indicate that noncompletion of treatment is associated with lower reductions in general and violent recidivism [12]. Furthermore, homework and registration assignments—an important part of all CBT programs—are often not completed in forensic outpatient populations. This negatively affects the transfer of psychological interventions from the therapist room to daily practice [15]. Second, the focus on cognitive control over emotional processes—a core element of most current treatment programs—may not fit the capacities of the patients. Forensic psychiatric patients often lack insight into their emotions [16] and have difficulty in observing and interpreting physiological signs of increased inner tension, such as accelerating heartbeat, sweating, or trembling. This limited capability to timely detect (particularly slowly) rising arousal levels increases the chance of *suddenly* occurring aggressive outbursts. Furthermore, cognitive control processes can be overruled by impulsive aggressive behavior in the case of high physiological arousal [17]. Therefore, in line with previous literature [18], therapy should focus more on strengthening awareness of bodily sensations associated with anger, before moving to the enhancement of cognitive control and deliberate responses in anger-provoking situations.

New technological applications may help address some of these treatment challenges. Technology-based interventions, such as mobile biofeedback apps [19], serious gaming [20], and virtual reality therapy [21] have the potential to increase adherence to treatment by engaging patients and by increasing maintenance during out-of-session activities [22]. Considering their lack of emotional awareness, interventions that provide the patients with information about their physiological state in real-life situations may help signal heightened arousal in response to emotional events and support adequate self-regulation [23,24]. In this study, we aim to explore the potential of a new biocueing intervention, which signals *at risk* levels of arousal in everyday life, as an addition to current aggression regulation therapy.

Biocueing can be seen as a specific, personalized type of biofeedback [25]. In the process of biofeedback, instruments monitor physiological parameters (eg, heart rate, skin conductance, and respiration), transform these measurements into auditory or visual signals, and present these signals to the user directly [26]. During a traditional, nonwearable biofeedback paradigm consisting of multiple on-site sessions, patients are

trained to regulate their physiological reactions by consciously alternating their responses to the given feedback. Rapid developments in noninvasive, wearable technology (eg, breast bands, wrist sensors, smart fibers, and interactive textiles) have opened opportunities for biocueing, which combines real-time measurement in everyday life and just-in-time behavioral support [27-29].

Traditional biofeedback has proven effective for patients with different psychopathology; however, most studies have been conducted among patients with internalizing problems [24]. Biocueing is relatively new but might be particularly useful for patients who lack insight into the physiological signals that precede dysregulated behavior in everyday life, such as binge-eating episodes [30] or self-injurious behavior [31]. Focusing on aggression, several pilot studies have provided the first evidence that physiological information can be used as a predictor of aggressive behavior, for example, among youth with autism spectrum disorder [32], patients with intellectual disabilities [33], and forensic patients [34]. Therefore, biocueing might be a helpful tool to increase awareness of high-risk situations and to support patients in practicing behavioral skills that prevent their escalation into aggressive incidents [35]. However, there is a gap between these study results and the actual deployment of mobile health (mHealth) interventions, such as a biocueing app for wearables, in forensic clinical practice. To bridge this gap, consideration of the needs of the intended users, as well as usability evaluation in the user's natural environment, is required [36]. The involvement of end

users throughout the design process, the core principle of user-centered design, is therefore highly recommended for the development of useful and effective mHealth interventions [37].

Study Aims

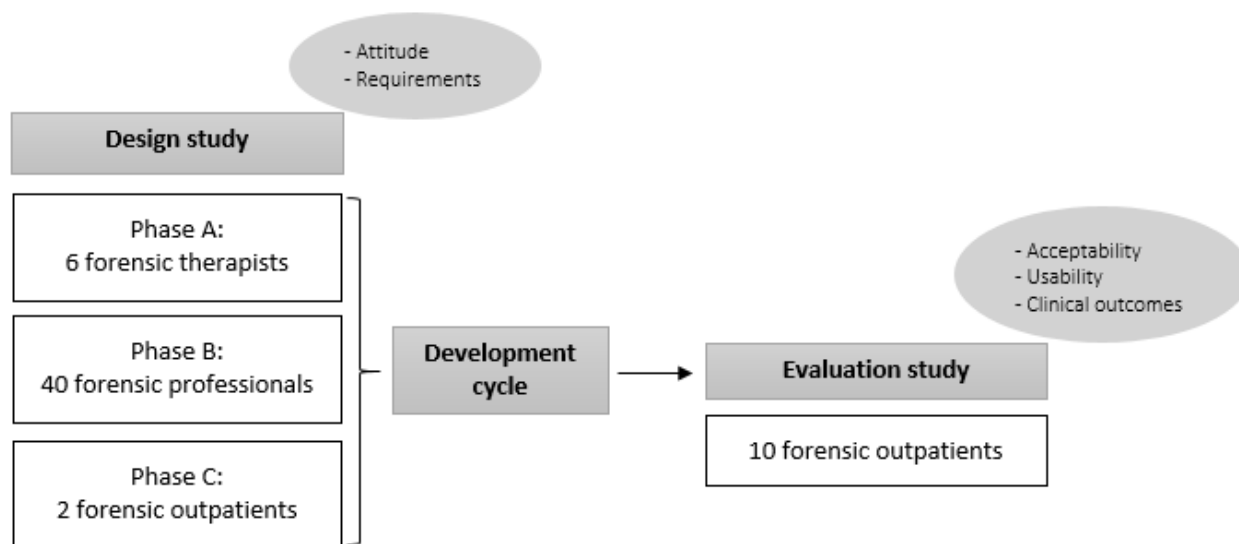
Given the potential of biocueing in dealing with the challenges in forensic (outpatient) treatment programs and using the principles of user-centered design, we first explored the attitudes of forensic professionals and patients toward this new intervention as an addition to aggression regulation therapy. In addition, we collected requirements to develop an updated version of the Sense-IT biocueing app [38] for use in a forensic outpatient population. Finally, we investigated the acceptability and usability of the revised Sense-IT app and explored changes related to aggression, anger, and interoceptive awareness in a 2-week evaluation study among 10 forensic outpatients with aggressive behavior.

Methods

Overview

First, we conducted a design study to explore the attitudes toward and specific requirements for a biocueing intervention in a forensic sample. With this information, an updated version of the Sense-IT app was developed. Then, we studied the app's acceptability and usability and its preliminary effects on clinical outcomes in a 2-week evaluation study. The structure of the study has been shown in [Figure 1](#).

Figure 1. Study design: phases, participants, and concepts.



Design Study

From October 2016 to March 2017, 6 forensic therapists, 40 forensic professionals (ie, therapists, psychiatric nurses, social workers, and probation officers), and 2 forensic outpatients participated in our study. We informed all the participants of the study's content and the voluntary basis for participation. We recruited the forensic therapists by email and the forensic professionals through in-person engagement at a forensic congress. The forensic outpatients were approached after

consultation with their therapists, and their signed informed consent was obtained. The forensic therapists and outpatients tested a precursor version of the Sense-IT app for 3 to 7 days, respectively, using a research-owned smartphone and smartwatch. The Sense-IT system provides a visual display of physiological arousal by measuring heart rate, notifies the user of level changes, and delivers a default message when the user's physiological arousal is significantly elevated ($SD > 2$) than their personal baseline. In this design study, the baseline measurement consisted of 300 reliable heart rate measures, with

20 seconds between each measurement. During the baseline measurement, participants could behave normally, without any restrictions. The forensic professionals responded to a short paper-and-pencil survey to increase their knowledge regarding their attitude toward bioeueing. None of the participants received any financial compensation for this study.

Given the outcomes of this study, the forensic professionals' attitudes toward bioeueing could be considered open and cautiously positive. The participants mentioned awareness of bodily signals accompanying anger (9/40, 22%) and insight into increasing arousal levels, especially in high-risk situations (15/40, 37%), as the most valuable additions to aggression regulation therapy. Furthermore, bioeueing was seen as a promising way to open therapeutic conversations about aggressive behavior. According to most forensic therapists (4/6, 67%), the usability of this precursor version of the Sense-IT app was insufficient. The limited battery life of the watch, the sudden watch face changes, the synchronization problems, the questionable reliability of heart rate measurements, the feedback method (too soft and too slow), and the mandatory use of a research phone were listed as items for improvement. Therapists were most satisfied with the visualization of arousal on the smartwatch and the warning function for heightened arousal. Of the 2 forensic outpatients, only one used the app for the entire week. This participant recommended audio-recording because he had difficulty typing notes. He reported an increased awareness of arousal, increased control over his aggressive behavior by the initiation of self-calming strategies, and distraction from inner tension by using the Sense-IT system. The other participant quit the study early because of technological shortcomings in this version of Sense-IT. This participant was frustrated with the app interrupting measurement when the watch face was accidentally touched. Given the insufficient usability scores and the increased irritability reported by one of the patients, we initiated a development cycle to resolve these technological shortcomings before further rolling out the app among forensic outpatients.

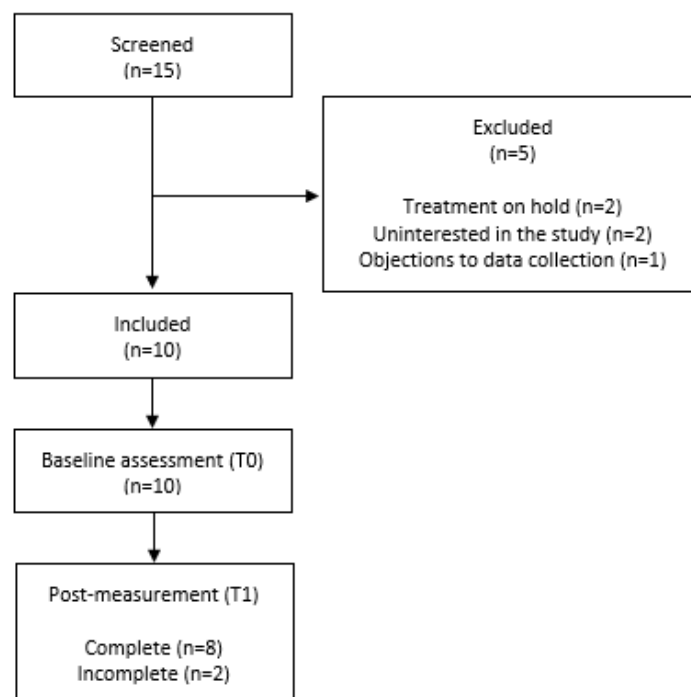
Development Cycle

Technological stabilization was the most important aim of this development cycle. The recommended improvements were implemented and intermediately tested by the researchers and app developers in 5 iteration rounds between 2017 and 2018. The activation of the Sense-IT app on the smartphone and smartwatch was synchronized and automated to prevent synchronization problems. Furthermore, the connection between the smartwatch and smartphone was made visible on the main screen. A clear on-off slider was incorporated into the main screen. Continuous visualization of the data on the watch face was optimized. Finally, the user was allowed to define during which activity profiles (eg, driving, cycling, and running) the operation of the Sense-IT should be paused. A description of the revised Sense-IT app has been provided in the next section.

Evaluation Study

Participants

We recruited forensic outpatients receiving aggression regulation therapy at Inforsa for participation from November 2018 to July 2019. Inforsa is a forensic mental health care facility that specializes in the treatment of patients with disruptive and criminal behavior. A research associate screened the potential participants for eligibility, in consultation with the patient's therapist. The eligibility criteria included (1) a proven lack of anger management skills, indicated by either a recently committed violent crime and/or a high risk of committing one, (2) assignment to individual outpatient aggression regulation treatment after multidisciplinary consultation, (3) basic understanding of mobile apps, and (4) aged ≥ 16 years. The exclusion criteria included (1) acute manic or psychotic symptoms, (2) current high risk of suicide, (3) severe addiction problems or other severe conditions requiring immediate intervention or hospitalization, and (4) insufficient understanding of the Dutch language. The first 3 exclusion criteria were assessed using cut-off scores on the corresponding items in the Health of the Nation Outcome Scales [39]. A total of 10 forensic outpatients were enrolled in the study. An outline of the recruitment and participation flow is displayed in [Figure 2](#).

Figure 2. Flow chart of recruitment and participation in the evaluation study.

Procedure

This study was approved by the Medical Ethical Committee of Amsterdam University Medical Centre, the Netherlands (NL63911.029.17). This study was also registered in the Netherlands Trial Register (NL8206). If a patient was eligible and interested in the research project, study participation was offered during a face-to-face appointment with the therapist and the patient. The research associate provided the patient with a brief oral description and full written information on the study. The voluntary nature and the absence of any negative consequences for the patient's refusal to participate were emphasized. If the patient was interested, the next appointment was planned after at least 7 days, providing enough time for consideration. In this appointment, we obtained the informed consent of the participants, and they filled out self-reported questionnaires. The baseline measurement (T0) lasted approximately 45 minutes. After completion of the assessment, participants were provided with a smartwatch and a mobile phone with the Sense-IT app. The participants were shown how to use the devices and were given tips on charging and using the system safely. They also received a user manual. The participants used the devices independently during the following 2 weeks. They were encouraged to call the research associates if any problems occurred. After the 2-week intervention period, another 45-minute assessment (T1) was planned. We used the same measurements as at T0, supplemented with qualitative interviews and quantitative usability measures.

Materials

Demographics

We collected demographic and clinical information at T0 using a 23-item self-developed questionnaire. The variables assessed included gender, ethnicity, judicial history, care history, education, family background, and social situation.

Mobile Phones and Smartwatches

The participants received both a smartwatch and a mobile phone. In this evaluation study, we used the Ticwatch E (Mobvoi, Ltd), a new smartwatch that had good reviews on reliability and cost-effectiveness and had a longer battery life than the smartwatches we used in the design study (Moto 360 2nd Gen; Lenovo Group, Ltd). The smartwatches were equipped with a photoplethysmography (PPG) sensor, by which the blood volume pulse can be measured and the heart rate can be derived. Connection with the mobile phone, the Moto C Plus (with Android 8.0 operating system; Google, LLC), was established via Bluetooth. We provided the participants with research-owned mobile phones to maintain control over the app settings and to account for secure data extraction.

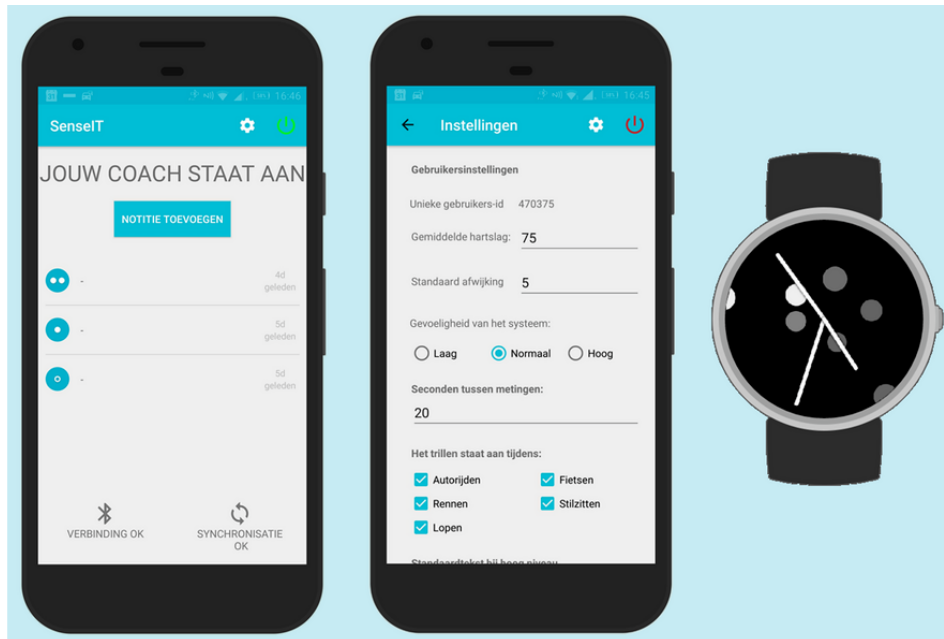
Sense-IT App

The newly developed version of the Sense-IT app, version 2.13, was preinstalled on all smartwatches and mobile phones before distribution. The Sense-IT app was originally developed by researchers at the University of Twente in co-operation with Scelta, an expert center for psychiatric patients with personality disorders [38,40]. The Sense-IT system reads the physiological data measured by the photoplethysmography (PPG) sensor and stores the data in a local database on the smartphone itself. The built-in algorithm compares the current heart rate of a user with their mean heart rate at baseline and calculates a level between -3 and 5 using the SD of the baseline measurement. In this study, the determination of the user's baseline was started at the end of the T0 measurement and lasted until the PPG sensor received 200 reliable heart rate measures, with 20 seconds between each measurement. During baseline measurement, participants could behave as they normally would, except for any intense physical activity. After this measurement, the level of their current heart rate was visually displayed on the smartwatch and was changed when the heart rate decreased or

increased by ≥ 1 SD. For this study, notifying vibrations were sent to the users at every change in their physiological level. The Sense-IT app detects (physical) activity categories using the accelerometer and Google activity recognition algorithms, allowing the user to receive notifications for certain activity profiles (eg, sitting still and driving a car) only. From the user interface on the smartphone, users can turn the app on and off, opening a timeline of all the measurement events and level changes detected by the system. Users can add notifications to events in the timeline and report their subjective level of arousal,

which might be particularly useful when the user is notified of level changes. Users can also define a personalized message that is displayed when their physiological arousal exceeds a predefined level. In this study, we used a default message (ie, “your heart rate is higher than average”), which was only displayed at levels 3, 4, and 5 ($SD > 2$) above baseline. The app’s user interface also presents information about the connection and synchronization status, as well as a settings page protected by a password to prevent unwarranted changes. Screenshots of the Sense-IT app are displayed in Figure 3.

Figure 3. Screenshots of the Sense-IT app (version 2.13): main screen with measurements, settings screen, and the watch face.



Acceptability

In a semistructured qualitative interview developed by the study team, the participants’ attitude toward technological interventions and their perceived proficiency in using new technology were assessed using a 5- and 10-point Likert scale, respectively. Closed-ended and open-ended questions assessed whether the participants would use the Sense-IT app in the future and whether they expected others to do so. The total number of heart rate measurements, measured every 20 seconds by the PPG sensor of the smartwatch, was used as an indicator of the actual use of the Sense-IT system. The damage, loss, and theft of the devices were recorded.

Usability

We administered the System Usability Scale (SUS), a short, commonly used questionnaire for quick and reliable assessment of product usability [41], at T1. The SUS consists of 10 statements that can be scored on a 5-point Likert scale, ranging from 1 (*totally disagree*) to 5 (*totally agree*). The SUS yields an overall score between 0 and 100, with higher scores indicating better usability. In the original study, 68 was used as a cut-off score. According to more recent research [42], a product is *acceptable* with scores above 70; better products score in the high 70s to upper 80s and superior products score above 90. Products with scores lower than 70 should be considered as candidates for improvement.

Furthermore, we evaluated usability qualitatively by using semistructured interviews. The interview included Likert scale questions about the attractiveness of the devices, the ease-of-use of the app, the clarity of watch faces, and the evaluation of feedback notifications. Open questions assessed the advantages and disadvantages of Sense-IT and any recommendations for its further improvement.

Aggression

We assessed for aggressive behavior using the Dutch version [43] of the Aggression Questionnaire-Short Form (AQ-SF) [44]. The AQ-SF is a self-report questionnaire, in which participants respond to 12 statements regarding aggression on a 5-point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The AQ-SF distinguishes 4 subscales: physical aggression, verbal aggression, anger, and hostility. The internal consistency coefficients for the total score of the Dutch AQ-SF ranged between 0.72 and 0.88 in a forensic population. Significant test-retest correlations after 4 weeks were found for the AQ-SF total and subscale scores, except for the physical aggression subscale. The AQ-SF was administered at T0 and T1.

Anger

We assessed anger and its subcomponents using the Dutch version [45] of the State-Trait Anger Expression Inventory-2 (STAXI-2) [46]. The STAXI-2 is a 57-item self-report

questionnaire, in which items are coded on a 4-point Likert scale ranging from 1 (*almost never*) to 4 (*almost always*). The questionnaire consists of 3 main scales: state anger, trait anger, and anger expression and anger control. The internal consistency, assessed in an inmate sample, was considered good, with Cronbach α ranging from .79 to .88. For the original version of the STAXI, the test-retest coefficients were acceptable, except for the State Anger scale. We administered the full STAXI-2 at T0; at T1, we administered only the state and trait anger scales because of time constraints.

Bodily Sensations Related to Anger

We measured bodily sensations related to anger in interpersonal situations using the Dutch version of a recently developed self-reported questionnaire called the Anger Bodily Sensations Questionnaire (ABSQ) [47]. In this 18-item self-report questionnaire, participants can rate their experience of physiological responses during anger-provoking interpersonal situations on a Likert scale, ranging from 1 (*not at all*) to 5 (*very much*). The internal consistency (Cronbach α) for the total score was .93 in an offender population for this study. The total score had good 1-week test-retest reliability within the offender sample. The ABSQ was administered at T0 and T1.

Data Analysis

We analyzed the quantitative data using SPSS (version 25, IBM Corp). Although this evaluation study was not intended for

inferential statistics, we exploratively compared the pre-post scores, after checking the normality assumptions, with the nonparametric equivalent of the paired *t* test, the Wilcoxon Matched Pairs test. We analyzed the qualitative data using Microsoft Word and Excel. Textual responses were first inspected for theme analysis, then coded into categories and described as relative results. Categorical responses were described as relative results.

Results

Demographics

Of the 10 forensic outpatients participating in this study, the majority (9/10, 90%) were male, in line with the usual male-female distribution in forensic populations. Although all participants were born in the Netherlands, out of 10 participants, 6 (60%) had parents originating from another country. Furthermore, 60% (6/10) reported problems such as domestic violence, substance abuse, or psychological problems in the families they grew up in. Most of the participants were referred to Inforsa for mandatory treatment as part of a conditional sentence and had been convicted multiple times in the past. The descriptive characteristics of the participants have been summarized in Table 1.

Table 1. Summary of demographic characteristics of the participants^a (N=10).

Demographic characteristics	Values
Age (years), mean (SD)	34.90 (13.29)
Gender, n (%)	
Male	9 (90)
Female	1 (10)
Educational background, n (%)	
Primary school	2 (20)
Secondary school	2 (20)
Secondary vocational education	6 (60)
Living conditions, n (%)	
Private home	4 (40)
Assisted living facility	4 (40)
Social care	2 (20)
Index offense, n (%)	
Violent crime	5 (50)
Gun crime	1 (10)
No index offense	4 (40)
Multiple convictions, n (%)	8 (80)
Mandatory treatment, n (%)	6 (60)
Previous ART ^b treatment, n (%)	5 (50)

^aNumber of participants measured at T0.

^bART: aggression replacement training.

Acceptability

The participants were provided with the Sense-IT system for 2 weeks. All participants returned these research-owned smartwatches and mobile phones. Although we noticed some superficial user damage and had to buy new charging cables, none of the devices had to be replaced because of damage, loss, or theft. Half of the participants (5/10, 50%) fully agreed with the statement that they liked to use technology or technological gadgets, a substantial number (4/10, 40%) claimed a neutral position and 1 participant (1/10, 10%) disagreed with the statement. On a 10-point Likert scale, the participants rated themselves as proficient in using new technologies (mean 7.1, SD 2.5 participants). At T1, it appeared that one participant had not used the Sense-IT app at all; the T1-responses of this participant were therefore excluded from further analysis. Among the other participants, the total amount of heart rate measurements (in hours) was used to indicate actual app use. As some participants showed very large numbers of measurements per day, we corrected the data for very low heart rates (<50 beats per minute) and measurements during nighttime (from midnight until 6 AM). The corrected actual app usage strongly differed between participants (mean 62.19, SD 38.63 hours; range 3.49-127.02 hours). Higher heart rate measurements were found among older participants, $r_8=-.768$, $P=.016$. No significant correlations were found among the attitude toward new technology, the perceived proficiency, and the indicator of actual app use. Furthermore, out of 9 participants at T1, 6 (67%) reported that they would like to use the Sense-IT app in the future. All the participants expected others to use the app in the future.

Usability

Participants who said they would not use the Sense-IT app in the future reported that the app had no added value for them because, for example, they did not regard themselves as aggressive or they already claimed to know their personal precursors for aggressive behavior. One participant therefore recommended the addition of the Sense-IT app in the early phases of treatment. The participants who would use the app in the future listed several conditions that could be considered as recommendations to further improve the Sense-IT app. The

average score on the SUS for the group was above the cut-off value (mean 73.1, SD 16.2). Significantly higher system usability scores were reported by participants with a more positive attitude toward new technology, $r_9=0.857$, $P=.002$; no significant correlations were found between the usability scores and perceived proficiency in using new technology. Most participants (8/9, 89%) did not report difficulty using the Sense-IT app on the smartphone. The watch faces on the smartwatch were reported to be clearly visible by the participants. The design of the Sense-IT app on the smartphone was considered neutral by half of the participants and (quite to very) attractive by the other half; some participants reported that they would like a more colorful design. Considering the messages shown when the physiological values exceeded a predefined level, out of 9 participants, 6 (67%) said they would like to use a default text message; the other participants (3/9, 33%) preferred a personalized message. The number of notifying vibrations (delivered at every level change) was considered too large by 56% (5/9) of the participants and was therefore most often mentioned as a point of improvement. Furthermore, improved accuracy, longer smartwatch battery life, and the possibility of using the Sense-IT app on their own smartphones were recommended.

Clinical Outcomes

We performed an exploratory analysis on the clinical outcome measures. Given the small sample size, the Shapiro-Wilk test was used to evaluate for normality of the data. As expected, the normality assumption was not met for several subscales of the measures used. Therefore, the Wilcoxon Matched Pairs test was used. No significant changes were found on the AQ-SF and ABSQ. Between baseline (median 2.35) and postmeasurement (median 1.90), trait anger measured with the STAXI-2 decreased significantly ($Z=-2388$; $P=.017$). Explorative visual analysis of the data showed that the scores of most of the participants (7/9, 78%) decreased. None of the participants showed elevated scores at postmeasurement compared with baseline. No significant correlations were found between this trait anger decrease and other variables such as attitude toward new technology, perceived proficiency in using new technology, and usability of the Sense-IT app. All clinical outcomes are shown in [Table 2](#).

Table 2. Summary of the clinical outcomes.

Clinical outcome scores	Time points	
	T0 (n=10), mean (SD)	T1 (n=9), mean (SD)
Aggression (AQ-SF^a)	2.36 (0.78)	2.41 (0.85)
Physical aggression	2.23 (1.21)	2.23 (1.32)
Verbal aggression	2.23 (0.96)	2.63 (0.78)
Anger	2.87 (1.00)	2.73 (1.31)
Hostility	2.10 (1.03)	2.03 (1.08)
Anger (STAXI-2^b)	1.91 (0.35)	— ^c
State anger	1.05 (0.13)	1.02 (0.05)
Trait anger	2.21 (0.65)	1.98 (0.72)
Anger expression and control	2.22 (0.48)	—
Bodily sensations related to anger (ABSQ ^d)	2.34 (0.93)	2.49 (0.95)

^aAQ-SF: Aggression Questionnaire-Short Form.

^bSTAXI-2: State-Trait Anger Expression Inventory-2.

^cMissing data.

^dABSQ: Anger Bodily Sensations Questionnaire.

Discussion

Principal Findings

To our knowledge, this study is one of the first attempts to enroll and evaluate a smartwatch-based biocueing intervention in a forensic outpatient population with aggression regulation difficulties [25]. Our study revealed a cautiously positive attitude toward the use of biocueing as an addition to regular therapy. Requirements for improvement were processed in a development cycle, resulting in an updated version of the Sense-IT app. The results of our 2-week evaluation study showed adequate usability scores, although the actual use of the app and its expected future use did not entirely match these outcomes. Furthermore, a significant decrease in trait anger was observed postintervention. Valuable recommendations were obtained for further improvement in Sense-IT. Considering the aim of our study, we were able to collect relevant information for the further development and enrollment of technology-based interventions as adjuncts to treatment, even in populations with lower treatment adherence.

In accordance with the principles of user-centered design for mHealth applications [37], which had also been applied in the earlier development phases of the Sense-IT app [38,40], we involved patients, therapists, and other forensic professionals in the development process of Sense-IT. The recommendations collected in the design study were mainly related to technological issues impeding the ease-of-use, such as synchronization problems and limitations in the battery life of the smartwatch. As the usability of the precursor version of the Sense-IT turned out to be inadequate for extensive testing among forensic outpatients, its further development was initiated. To generate more input for this development cycle, the professionals' attitudes toward the use of a biocueing app for this particular group were more thoroughly assessed. Our design study indicated that forensic professionals had an open and

positive attitude toward biocueing, recognizing the potential disadvantages or risks. This is important because openness toward new treatment possibilities and a belief that the intervention might be beneficial to patients are, among others, considered as facilitators of the use of mHealth technology [48,49].

Patients' adoption of mHealth applications is also known to be a result of several factors, such as perceived usefulness and ease-of-use, influencing their individual attitude and behavioral intention, mediated by age [50]. Other results suggest that the perceived mobile technology identity, related information technology experience, and self-efficacy are associated with higher adoption rates [51]. In our evaluation study, the users' perceived proficiency was not associated with the actual use and the perceived usability of Sense-IT. Actual use seemed higher among older participants, contrary to the mediating influence of age, as mentioned in other studies. However, the higher usability scores being reported by patients with a more positive attitude toward new technology was consistent with these earlier findings. Furthermore, we encountered some skepticism when we first presented our idea to investigate a biocueing intervention among forensic outpatients by providing them with a smartwatch and smartphone. Contrary to expectations, none of the devices had to be replaced because of loss, damage, or theft, supporting a recent study among homeless youth [52]. The total amount of heart rate measurements, used as an indicator for the actual app usage, varied widely among the participants. Most participants used the Sense-IT quite often, but a few participants showed very low adherence to the app. This limited adherence to technology-based interventions is a common issue in eHealth and mHealth studies, even in samples that are more open to treatment. Many users of self-help applications show inconsistent use patterns [53], do not continue their use after completion of one exercise or module [54], or stop using a health app after 2 weeks, especially when their

preferences and goals are not met [55]. Given the answers regarding the future use of the Sense-IT app, actual usage of the app might have decreased even more if the duration of our study had been extended to more than 2 weeks. Nevertheless, considering the motivational problems and the high dropout rates in forensic populations, outpatients who often fail to practice outside the treatment setting would be encouraged to reflect on their behavior in real life.

Furthermore, the Sense-IT app received acceptable system usability scores, which means that the ease-of-use was considered *good enough* for further development and can be rolled out in this population. The participants provided valuable recommendations for a new development cycle. The number of notifying vibrations, which were considered disturbing by most of the participants, was most often mentioned as a point for improvement. For some patients, this feature of the app led to increased irritability, a possibility that has already been mentioned in previous literature [56-58]. In the future, this drawback can easily be remedied by adjusting the levels at which notifying vibrations are provided and by allowing users to customize the settings themselves. Furthermore, our study revealed that some participants had difficulty coming up with a personalized message, which could be shown when their physiological values were elevated. Therefore, their preference for the default text message might have been a choice out of convenience. Further integration of the Sense-IT app in therapy may help to overcome this difficulty. At the design level, desired improvements, such as the addition of multiple colors, were mostly mentioned by younger participants. Furthermore, several participants would have liked to use Sense-IT on their own mobile phones. For clinical use, outside a research context with privacy constraints, this would certainly be possible and might facilitate the acceptability and usability of the app. Finally, technological hardware improvements, such as the extended battery life of the smartwatch, could further enhance adherence to Sense-IT.

To investigate whether the 2 weeks of using the Sense-IT biocueing app were associated with clinically relevant changes, we performed exploratory analysis on the pre and post measurement outcomes of aggression, anger, and anger bodily sensations. In line with the literature indicating the difficulty of changing aggressive behavior [13] and given the small sample size and short intervention period of our study, no significant changes were found in the overall scores. Considering the concepts of state and trait anger, in this study, change might theoretically occur on state anger. Interestingly, trait anger diminished significantly between pre- and postintervention measurements. No change was found in the state anger. However, this might have been affected by the fact that state anger was only measured at 2 specific moments in time, when the participants were generally in a resting state. Regarding the change in trait anger, social desirability might offer a partial explanation for this finding. In a study among forensic inpatients, participants with high scores on impression management reported significantly lower levels of trait anger [59]. To explore whether this finding reflects a true change in the frequency of experiencing angry feelings should therefore be more thoroughly assessed within a longitudinal research

design, using assessment methods that are more sensitive to minor changes and less susceptible to social desirability.

Limitations

Our study had several limitations. First, the number of forensic outpatients participating in the design and evaluation study was low. In line with the early phase of biocueing research and the often-encountered difficulties in studies among forensic patients, this may have affected the findings of our study. Second, we used a small subset of questionnaires in our evaluation study, focusing on anger and aggression. Therefore, we might have overlooked other relevant changes, for example, in emotion regulation in general. In addition, we did not use the full STAXI-2 at T1, which should have been preferred to prevent data loss. Furthermore, the questionnaires we used might have been susceptible to social desirability. In any case, the most used questionnaires to assess aggressive behavior are not designed to detect small changes over short periods. In this regard, the short duration of the intervention and the absence of a follow-up measurement complicate the interpretation of our findings. Third, because we used a pretest-posttest design without a control group, we were not able to disentangle the impact of the use of the Sense-IT biocueing intervention from aggression regulation therapy and from the distribution of mobile phones and smartwatches, which might be considered an intervention itself. Fourth, in some cases, Sense-IT seemed to be measuring data while not being actually worn by the participants. This complicated the interpretation of the actual usage of the device. Fifth, as Sense-IT could not be used on the participants' own phones, owing to privacy constraints in research, it might have had a restrictive effect on adherence. Sixth, the explanation and handing out of the devices was done by research associates in the absence of their therapists. In addition, both patients and therapists were not able to adjust the settings during the study. These restrictions, associated with the research design, might have limited the adoption of the Sense-IT app by therapists, an important driver for the integration of mHealth interventions in clinical practice [49].

Implications for Future Research and Practice

This development and usability study has several key implications for future research. First, it emphasized the importance of the involvement of both patients and therapists in the development of effective mHealth interventions [35]. Second, the evaluation study yielded new recommendations for the improvement of the Sense-IT app. At the technological level, stabilization of the app should remain a critical area for improvement. The number of notifying vibrations and the levels at which SMS text messages are sent should be adjustable to the wishes and needs of patients. In addition, the ease-of-use of recording subjective arousal levels should be enhanced. At the design level, the clarity of the main screen should be improved, and the measurement screen should be updated with a semigraphical representation. To further enhance the attractiveness of Sense-IT, some patients would like to choose different background colors. Meanwhile, the recommendations were processed in a new development cycle, resulting in an updated version of Sense-IT, as shown in Figure 4. Changes in app settings and the use of new smartwatches should be

considered to improve the battery life and, thereby, usability. Third, experimental studies should be performed to evaluate the effectiveness of the Sense-IT intervention on clinical outcomes. Single-case experimental designs should be considered as well because these types of designs might be better able to detect minor changes over short periods by, for example, using ecological momentary assessment [60,61]. Laboratory tasks could also be considered to gain a broader insight into the response to bio cueing. Fourth, the integration of therapy should be enhanced. To experience the potential benefits of this bio cueing intervention, therapists should first

be trained in working with Sense-IT and familiarizing themselves with the app. Furthermore, both therapists and patients should be able to adjust the settings to explore the optimal level of sensitivity of the system. Fifth, the Sense-IT app should preferably be added to treatment in the early phases of aggression regulation therapy, focusing on the recognition of anger bodily signals. The use of new technological interventions might have a positive influence on treatment motivation [21], especially among patients with a positive attitude toward new technology.

Figure 4. Screenshots of the Sense-IT app after further development (version 2.57): main screen, measurement screen, note screen, and the watch face.



Conclusions

This study revealed a cautiously positive attitude toward the use of bio cueing as an addition to regular aggression regulation therapy in forensic psychiatry. In the evaluation study among forensic outpatients, the revised version of the Sense-IT app demonstrated moderate acceptability and adequate usability. Furthermore, a significant decrease in trait anger was found postintervention, which should be further explored in future research using appropriate research designs. Valuable recommendations for improvement of Sense-IT, both at the technological and design levels, were obtained in this study. For patients and therapists, a stable functioning app, with few synchronization disruptions and a self-adjustable number of notifications, seemed most important. Implementing new

smartwatches with a longer battery life and using Sense-IT on the user's own smartphone are expected to increase adherence in the future. Considering the actual use of Sense-IT, the app seemed to have facilitated out-of-session practice and might therefore represent an alternative for more traditional paper-and-pencil registration assignments. The extent to which users actually reflect on their behavior and whether they feel supported by Sense-IT to practice behavioral alternatives needs further examination. Furthermore, this study provided some evidence that the deployment of Sense-IT is most useful in the first phase of aggression regulation therapy. Finally, to enhance the app's integration in treatment, therapists should be trained in the use of the app to facilitate exploration of the potential benefits of these kinds of new mHealth interventions with their patients.

Acknowledgments

This study was performed within a larger research project financially supported by the Dutch Ministry of Justice and Security. The authors would like to thank the study participants and the therapists at Inforsa, forensic psychiatric care in Amsterdam, the Netherlands. Furthermore, the authors thank Eline Middelhoven, Christiaan Doll, Bo Hornkamp, Lotte Nijhuis, Michelle Kemper, and Nienke Smit for their assistance in data collection. Finally, they would like to thank the organizations involved in the development of Sense-IT for use in this study project: University of Twente, Scelta/GGNet, Amsterdam University Medical Center, Inforsa/Arkin, and Plury.

Authors' Contributions

AtH, TvdP, LS, AP, and MN conceived the study. AtH, TvdP, LS, AP, and MLN developed the study design. AtH and MLN tested versions of the Sense-IT app during the development cycle and provided the app developers with feedback. AtH coordinated the data collection and data processing. TvdP, MLN, and AEG helped AtH interpret the data. AtH drafted the manuscript in close co-operation with MLN. The other authors have critically revised the manuscript. All the authors have approved the final content of this manuscript.

Conflicts of Interest

AtH and MLN (and more remotely TvdP and AP) were involved in the development process of the Sense-IT app.

References

1. Krug EG, Mercy JA, Dahlberg LL, Zwi AB. The world report on violence and health. *Lancet* 2002 Oct 05;360(9339):1083-1088. [doi: [10.1016/S0140-6736\(02\)11133-0](https://doi.org/10.1016/S0140-6736(02)11133-0)] [Medline: [12384003](https://pubmed.ncbi.nlm.nih.gov/12384003/)]
2. Bouffard LA, Koeppl MD. Understanding the potential long-term physical and mental health consequences of early experiences of victimization. *Justice Q* 2014;31(3):568-587. [doi: [10.1080/07418825.2012.734843](https://doi.org/10.1080/07418825.2012.734843)]
3. Steadman HJ, Osher FC, Robbins PC, Case B, Samuels S. Prevalence of serious mental illness among jail inmates. *Psychiatr Serv* 2009 Jun;60(6):761-765. [doi: [10.1176/ps.2009.60.6.761](https://doi.org/10.1176/ps.2009.60.6.761)] [Medline: [19487344](https://pubmed.ncbi.nlm.nih.gov/19487344/)]
4. Ogloff JR, Talevski D, Lemphers A, Wood M, Simmons M. Co-occurring mental illness, substance use disorders, and antisocial personality disorder among clients of forensic mental health services. *Psychiatr Rehabil J* 2015 Mar;38(1):16-23. [doi: [10.1037/prj0000088](https://doi.org/10.1037/prj0000088)] [Medline: [25799303](https://pubmed.ncbi.nlm.nih.gov/25799303/)]
5. The high budgetary cost of incarceration. Center for Economic Policy and Research. 2010. URL: <https://www.cepr.net/documents/publications/incarceration-2010-06.pdf> [accessed 2021-03-17]
6. Rubio-Valera M, Luciano JV, Ortiz JM, Salvador-Carulla L, Gracia A, Serrano-Blanco A. Health service use and costs associated with aggressiveness or agitation and containment in adult psychiatric care: a systematic review of the evidence. *BMC Psychiatry* 2015 Mar 04;15:35 [FREE Full text] [doi: [10.1186/s12888-015-0417-x](https://doi.org/10.1186/s12888-015-0417-x)] [Medline: [25881240](https://pubmed.ncbi.nlm.nih.gov/25881240/)]
7. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. *Behav Ther* 2016 Sep;47(5):755-772. [doi: [10.1016/j.beth.2016.05.002](https://doi.org/10.1016/j.beth.2016.05.002)] [Medline: [27816086](https://pubmed.ncbi.nlm.nih.gov/27816086/)]
8. Fairburn CG, Patel V. The impact of digital technology on psychological treatments and their dissemination. *Behav Res Ther* 2017 Jan;88:19-25 [FREE Full text] [doi: [10.1016/j.brat.2016.08.012](https://doi.org/10.1016/j.brat.2016.08.012)] [Medline: [28110672](https://pubmed.ncbi.nlm.nih.gov/28110672/)]
9. Glick B, Gibbs J. *Aggression Replacement Training: A Comprehensive Intervention for Aggressive Youth*. Champaign IL: Research Press; 2010.
10. Landenberger NA, Lipsey MW. The positive effects of cognitive-behavioral programs for offenders: a meta-analysis of factors associated with effective treatment. *J Exp Criminol* 2005 Dec 10;1(4):451-476. [doi: [10.1007/s11292-005-3541-7](https://doi.org/10.1007/s11292-005-3541-7)]
11. Lipsey MW, Landenberger NA, Wilson SJ. Effects of cognitive-behavioral programs for criminal offenders. *Campbell Syst Rev* 2007 Aug 13;3(1):1-27. [doi: [10.4073/csr.2007.6](https://doi.org/10.4073/csr.2007.6)]
12. Henwood KS, Chou S, Browne KD. A systematic review and meta-analysis on the effectiveness of CBT informed anger management. *Aggress Viol Behav* 2015;25:280-292. [doi: [10.1016/j.avb.2015.09.011](https://doi.org/10.1016/j.avb.2015.09.011)]
13. Brännström L, Kaunitz C, Andershed A, South S, Smedslund G. Aggression replacement training (ART) for reducing antisocial behavior in adolescents and adults: a systematic review. *Aggress Viol Behav* 2016;27:30-41. [doi: [10.1016/j.avb.2016.02.006](https://doi.org/10.1016/j.avb.2016.02.006)]
14. Olver ME, Stockdale KC, Wormith JS. A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *J Consult Clin Psychol* 2011 Feb;79(1):6-21. [doi: [10.1037/a0022200](https://doi.org/10.1037/a0022200)] [Medline: [21261430](https://pubmed.ncbi.nlm.nih.gov/21261430/)]
15. Fletcher R, Tam S, Omojola O, Redemski R, Kwan J. Wearable sensor platform and mobile application for use in cognitive behavioral therapy for drug addiction and PTSD. In: Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2011 Presented at: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 30 -Sep 3, 2011; Boston, MA, USA. [doi: [10.1109/iembs.2011.6090513](https://doi.org/10.1109/iembs.2011.6090513)]
16. Robertson T, Daffern M, Bucks RS. Beyond anger control: difficulty attending to emotions also predicts aggression in offenders. *Psychol Viol* 2015 Jan;5(1):74-83. [doi: [10.1037/a0037214](https://doi.org/10.1037/a0037214)]
17. Kahneman D. A perspective on judgment and choice: mapping bounded rationality. *Am Psychol* 2003 Sep;58(9):697-720. [doi: [10.1037/0003-066X.58.9.697](https://doi.org/10.1037/0003-066X.58.9.697)] [Medline: [14584987](https://pubmed.ncbi.nlm.nih.gov/14584987/)]
18. Novaco R. Anger dysregulation. In: *Anger, Aggression and Interventions for Interpersonal Violence*. USA: Lawrence Erlbaum Associates Publishers; 2007.
19. Mackintosh M, Niehaus J, Taft CT, Marx BP, Grubbs K, Morland LA. Using a mobile application in the treatment of dysregulated anger among veterans. *Mil Med* 2017 Nov;182(11):e1941-e1949. [doi: [10.7205/MILMED-D-17-00063](https://doi.org/10.7205/MILMED-D-17-00063)] [Medline: [29087863](https://pubmed.ncbi.nlm.nih.gov/29087863/)]
20. Smeijers D, Koole SL. Testing the effects of a Virtual Reality Game for Aggressive Impulse Management (VR-GAIME): study protocol. *Front Psychiatry* 2019 Feb 26;10:83 [FREE Full text] [doi: [10.3389/fpsyt.2019.00083](https://doi.org/10.3389/fpsyt.2019.00083)] [Medline: [30863328](https://pubmed.ncbi.nlm.nih.gov/30863328/)]

21. Klein Tunte S, Bogaerts S, Bulten E, Keulen-de Vos M, Vos M, Bokern H, et al. Virtual Reality Aggression Prevention Therapy (VRAPT) versus waiting list control for forensic psychiatric inpatients: a multicenter randomized controlled trial. *J Clin Med* 2020 Jul 16;9(7):2258 [FREE Full text] [doi: [10.3390/jcm9072258](https://doi.org/10.3390/jcm9072258)] [Medline: [32708637](https://pubmed.ncbi.nlm.nih.gov/32708637/)]
22. Reiner R. Integrating a portable biofeedback device into clinical practice for patients with anxiety disorders: results of a pilot study. *Appl Psychophysiol Biofeedback* 2008 Mar;33(1):55-61. [doi: [10.1007/s10484-007-9046-6](https://doi.org/10.1007/s10484-007-9046-6)] [Medline: [18286369](https://pubmed.ncbi.nlm.nih.gov/18286369/)]
23. Lisetti CL, Nasoz F. Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP J Adv Signal Process* 2004 Sep 18;2004(11):- [doi: [10.1155/s1110865704406192](https://doi.org/10.1155/s1110865704406192)]
24. Yucha C, Montgomery D. Evidence-Based Practice in Biofeedback and Neurofeedback. Wheat Ridge: Association for Applied Psychophysiology and Biofeedback; 2008:-67.
25. Ter Harmsel JF, Noordzij M, Goudriaan A, Dekker J, Swinkels L, van der Pol T, et al. Biocueing and ambulatory biofeedback to enhance emotion regulation: a review of studies investigating non-psychiatric and psychiatric populations. *Int J Psychophysiol* 2021 Jan;159:94-106 [FREE Full text] [doi: [10.1016/j.ijpsycho.2020.11.009](https://doi.org/10.1016/j.ijpsycho.2020.11.009)] [Medline: [33248196](https://pubmed.ncbi.nlm.nih.gov/33248196/)]
26. McKee MG. Biofeedback: an overview in the context of heart-brain medicine. *Cleve Clin J Med* 2008 Mar;75 Suppl 2:S31-S34 [FREE Full text] [doi: [10.3949/ccjm.75.suppl_2.s31](https://doi.org/10.3949/ccjm.75.suppl_2.s31)] [Medline: [18540143](https://pubmed.ncbi.nlm.nih.gov/18540143/)]
27. Riley WT, Serrano KJ, Nilsen W, Atienza AA. Mobile and wireless technologies in health behavior and the potential for intensively adaptive interventions. *Curr Opin Psychol* 2015 Oct 01;5:67-71 [FREE Full text] [doi: [10.1016/j.copsyc.2015.03.024](https://doi.org/10.1016/j.copsyc.2015.03.024)] [Medline: [26086033](https://pubmed.ncbi.nlm.nih.gov/26086033/)]
28. Nahum-Shani I, Hekler EB, Spruijt-Metz D. Building health behavior models to guide the development of just-in-time adaptive interventions: a pragmatic framework. *Health Psychol* 2015 Dec;34S:1209-1219 [FREE Full text] [doi: [10.1037/hea0000306](https://doi.org/10.1037/hea0000306)] [Medline: [26651462](https://pubmed.ncbi.nlm.nih.gov/26651462/)]
29. Wang H, Yang C, Chen K, Chueh K. Relationship between heart rate variability and aggressive behavior among patients with schizophrenia hospitalized in acute wards. *Perspect Psychiatr Care* 2020 Apr;56(2):321-329. [doi: [10.1111/ppc.12433](https://doi.org/10.1111/ppc.12433)] [Medline: [31410853](https://pubmed.ncbi.nlm.nih.gov/31410853/)]
30. Smartphones for smarter eating: Elucidating eating behaviors, stress, and heart rate variability. UC San Diego. URL: <https://escholarship.org/uc/item/2p7547rt> [accessed 2021-02-23]
31. Koenig J, Rinnowitz L, Parzer P, Resch F, Thayer JF, Kaess M. Resting cardiac function in adolescent non-suicidal self-injury: the impact of borderline personality disorder symptoms and psychosocial functioning. *Psychiatry Res* 2017 Feb;248:117-120. [doi: [10.1016/j.psychres.2016.12.024](https://doi.org/10.1016/j.psychres.2016.12.024)] [Medline: [28039803](https://pubmed.ncbi.nlm.nih.gov/28039803/)]
32. Cumpanasoiu D, Mazefsky C, Stedman A, Peura C, Tian P, Guo Y, et al. 5.12 Physiological biomarkers for prediction of imminent aggression in minimally verbal children with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry* 2017 Oct;56(10):S256-S257. [doi: [10.1016/j.jaac.2017.09.295](https://doi.org/10.1016/j.jaac.2017.09.295)]
33. Palix J, Akselrod M, Cungi C, Giuliani F, Favrod J. Changes in heart rate variability recorded in natural situation with t-shirt integrated sensors and level of observed behavioral excitation: a pilot study of patients with intellectual disabilities and psychiatric disorders. *Front Psychiatry* 2017 Feb 1;8:4 [FREE Full text] [doi: [10.3389/fpsy.2017.00004](https://doi.org/10.3389/fpsy.2017.00004)] [Medline: [28203208](https://pubmed.ncbi.nlm.nih.gov/28203208/)]
34. de Looft P, Noordzij ML, Moerbeek M, Nijman H, Didden R, Embregts P. Changes in heart rate and skin conductance in the 30 min preceding aggressive behavior. *Psychophysiology* 2019 Oct;56(10):e13420. [doi: [10.1111/psyp.13420](https://doi.org/10.1111/psyp.13420)] [Medline: [31184379](https://pubmed.ncbi.nlm.nih.gov/31184379/)]
35. Neurowetenschappelijke toepassingen in de jeugdstrafrechtketen. Onderzoek en beleid 318. URL: <https://repository.wodc.nl/handle/20.500.12832/1269> [accessed 2021-01-05]
36. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. *Biomed Instrum Technol* 2012;Suppl:49-56. [doi: [10.2345/0899-8205-46.s2.49](https://doi.org/10.2345/0899-8205-46.s2.49)] [Medline: [23039777](https://pubmed.ncbi.nlm.nih.gov/23039777/)]
37. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform* 2016 Apr;60:243-251 [FREE Full text] [doi: [10.1016/j.jbi.2016.02.002](https://doi.org/10.1016/j.jbi.2016.02.002)] [Medline: [26903153](https://pubmed.ncbi.nlm.nih.gov/26903153/)]
38. Derks YP, Klaassen R, Westerhof GJ, Bohlmeijer ET, Noordzij ML. Development of an ambulatory biofeedback app to enhance emotional awareness in patients with borderline personality disorder: multicycle usability testing study. *JMIR Mhealth Uhealth* 2019 Oct 15;7(10):e13479 [FREE Full text] [doi: [10.2196/13479](https://doi.org/10.2196/13479)] [Medline: [31617851](https://pubmed.ncbi.nlm.nih.gov/31617851/)]
39. Wing JK, Beevor AS, Curtis RH, Park SG, Hadden J, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. *Br J Psychiatry* 1998 Jan;172(1):11-18. [doi: [10.1192/bjp.172.1.11](https://doi.org/10.1192/bjp.172.1.11)] [Medline: [9534825](https://pubmed.ncbi.nlm.nih.gov/9534825/)]
40. De Visser T, Bohlmeijer ET, Noordzij ML, Derks YP. mHealth in Mental Health: how to efficiently and scientifically create an ambulatory biofeedback e-coaching app for patients with borderline personality disorder. *Int J Hum Factors Ergon* 2017;5(1):61.
41. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry. Boca Raton, Florida: CRC Press; 1996.
42. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Human Comput Interac* 2008 Jul 30;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]

43. Hornsveld RH, Muris P, Kraaimaat FW, Meesters C. Psychometric properties of the aggression questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *Assessment* 2009 Jun;16(2):181-192. [doi: [10.1177/1073191108325894](https://doi.org/10.1177/1073191108325894)] [Medline: [18852480](https://pubmed.ncbi.nlm.nih.gov/18852480/)]
44. Buss AH, Perry M. The aggression questionnaire. *J Personal Soc Psychol* 1992 Sep;63(3):452-459. [doi: [10.1037/0022-3514.63.3.452](https://doi.org/10.1037/0022-3514.63.3.452)]
45. Hovens J, Lievaart M, Rodenburg J. STAXI-2. Vragenlijst over boosheid. Hogrefe. 2014. URL: https://www.hogrefe.com/nl/shop/media/downloads/sample-reports/5704201_InkijkexemplaarHandleiding_Samplepages.pdf [accessed 2021-10-18]
46. Spielberger CD. STAXI-2: State-Trait Anger Expression Inventory-2. USA: Psychological Assessment Resources; 1999.
47. Zwets AJ, Hornsveld R, Kraaimaat FW, Kanters T, Muris P, van Marle H. The psychometric properties of the Anger Bodily Sensations Questionnaire (ABSQ). *J Forensic Psychiatry Psychol* 2014 Jun 12;25(4):432-450. [doi: [10.1080/14789949.2014.925957](https://doi.org/10.1080/14789949.2014.925957)]
48. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. *J Am Med Inform Assoc* 2016 Jan;23(1):212-220 [FREE Full text] [doi: [10.1093/jamia/ocv052](https://doi.org/10.1093/jamia/ocv052)] [Medline: [26078410](https://pubmed.ncbi.nlm.nih.gov/26078410/)]
49. Feijt MA, de Kort YA, Bongers IM, IJsselstein 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/)]
50. 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)]
51. Balapour A, Reyshav I, Sabherwal R, Azuri J. Mobile technology identity and self-efficacy: implications for the adoption of clinically supported mobile health apps. *Int J Inf Manag* 2019 Dec;49:58-68. [doi: [10.1016/j.ijinfomgt.2019.03.005](https://doi.org/10.1016/j.ijinfomgt.2019.03.005)]
52. Schueller SM, Glover AC, Rufa AK, Dowdle CL, Gross GD, Karnik NS, et al. A mobile phone-based intervention to improve mental health among homeless young adults: pilot feasibility trial. *JMIR Mhealth Uhealth* 2019 Jul 02;7(7):e12347 [FREE Full text] [doi: [10.2196/12347](https://doi.org/10.2196/12347)] [Medline: [31267980](https://pubmed.ncbi.nlm.nih.gov/31267980/)]
53. Nelson EC, Sools AM, Vollenbroek-Hutten MM, Verhagen T, Noordzij ML. Embodiment of wearable technology: qualitative longitudinal study. *JMIR Mhealth Uhealth* 2020 Nov 03;8(11):e16973 [FREE Full text] [doi: [10.2196/16973](https://doi.org/10.2196/16973)] [Medline: [33141093](https://pubmed.ncbi.nlm.nih.gov/33141093/)]
54. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 06;20(6):e199 [FREE Full text] [doi: [10.2196/jmir.9275](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
55. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* 2019 Feb;18(1):97-98 [FREE Full text] [doi: [10.1002/wps.20592](https://doi.org/10.1002/wps.20592)] [Medline: [30600619](https://pubmed.ncbi.nlm.nih.gov/30600619/)]
56. De 'zelfmetende' justitiabele. Een verkennend onderzoek naar technologische zelfmeetmethoden binnen de justitiële context. *Cahiers*. 2017. URL: <https://repository.wodc.nl/handle/20.500.12832/195> [accessed 2021-01-05]
57. Howell N, Devendorf L, Vega GT, Tian R, Ryokai K. Tensions of data-driven reflection: a case study of real-time emotional biosensing. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018 Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; Apr 21-26, 2018; Montreal QC Canada. [doi: [10.1145/3173574.3174005](https://doi.org/10.1145/3173574.3174005)]
58. Owens J, Cribb A. 'My Fitbit Thinks I Can Do Better!' Do health promoting wearable technologies support personal autonomy? *Philos Technol* 2017 Jun 6;32(1):23-38. [doi: [10.1007/s13347-017-0266-2](https://doi.org/10.1007/s13347-017-0266-2)]
59. McEwan T, Davis MR, MacKenzie R, Mullen PE. The effects of social desirability response bias on STAXI-2 profiles in a clinical forensic sample. *Br J Clin Psychol* 2009 Nov;48(Pt 4):431-436. [doi: [10.1348/014466509X454886](https://doi.org/10.1348/014466509X454886)] [Medline: [19555524](https://pubmed.ncbi.nlm.nih.gov/19555524/)]
60. Smith JD. Single-case experimental designs: a systematic review of published research and current standards. *Psychol Methods* 2012 Dec;17(4):510-550 [FREE Full text] [doi: [10.1037/a0029312](https://doi.org/10.1037/a0029312)] [Medline: [22845874](https://pubmed.ncbi.nlm.nih.gov/22845874/)]
61. Maric M, van der Werff V. Single-case experimental designs in clinical intervention research. In: *Small Sample Size Solutions*. Oxfordshire: Routledge; 2020.

Abbreviations

- ABSQ:** Anger Bodily Sensations Questionnaire
- AQ-SF:** Aggression Questionnaire-Short Form
- ART:** aggression replacement training
- CBT:** Cognitive Behavioral Therapy
- mHealth:** mobile health
- PPG:** photoplethysmography
- STAXI-2:** State-Trait Anger Expression Inventory-2
- SUS:** System Usability Scale

Edited by G Eysenbach; submitted 31.03.21; peer-reviewed by T Alessa, N Karnik; comments to author 25.04.21; revised version received 13.05.21; accepted 07.09.21; published 24.11.21.

Please cite as:

ter Harmsel A, van der Pol T, Swinkels L, Goudriaan AE, Popma A, Noordzij ML

Development of a Wearable Biocueing App (Sense-IT) Among Forensic Psychiatric Outpatients With Aggressive Behavior: Design and Evaluation Study

JMIR Form Res 2021;5(11):e29267

URL: <https://formative.jmir.org/2021/11/e29267>

doi: [10.2196/29267](https://doi.org/10.2196/29267)

PMID: [34821567](https://pubmed.ncbi.nlm.nih.gov/34821567/)

©Annemieke ter Harmsel, Thimo van der Pol, Lise Swinkels, Anna E Goudriaan, Arne Popma, Matthijs L Noordzij. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 24.11.2021. 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 Phenotyping Algorithm to Identify People With HIV in Electronic Health Record Data (HIV-Phen): Development and Evaluation Study

Sarah B May^{1,2,3}, MS, MPH; Thomas P Giordano^{2,3,4}, MD, MPH; Assaf Gottlieb¹, PhD

¹School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX, United States

²Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, TX, United States

³Center for Innovation in Quality, Effectiveness and Safety, Michael E DeBakey VA Medical Center, Houston, TX, United States

⁴Section of Infectious Diseases, Department of Medicine, Baylor College of Medicine, Houston, TX, United States

Corresponding Author:

Assaf Gottlieb, PhD

School of Biomedical Informatics

University of Texas Health Science Center at Houston

7000 Fannin, Suite 600

Houston, TX, 77030

United States

Phone: 1 713 500 3698

Email: Assaf.Gottlieb@uth.tmc.edu

Abstract

Background: Identification of people with HIV from electronic health record (EHR) data is an essential first step in the study of important HIV outcomes, such as risk assessment. This task has been historically performed via manual chart review, but the increased availability of large clinical data sets has led to the emergence of phenotyping algorithms to automate this process. Existing algorithms for identifying people with HIV rely on a combination of International Classification of Disease codes and laboratory tests or closely mimic clinical testing guidelines for HIV diagnosis. However, we found that existing algorithms in the literature missed a significant proportion of people with HIV in our data.

Objective: The aim of this study is to develop and evaluate HIV-Phen, an updated criteria-based HIV phenotyping algorithm.

Methods: We developed an algorithm using HIV-specific laboratory tests and medications and compared it with previously published algorithms in national and local data sets to identify cohorts of people with HIV. Cohort demographics were compared with those reported in the national and local surveillance data. Chart reviews were performed on a subsample of patients from the local database to calculate the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of the algorithm.

Results: Our new algorithm identified substantially more people with HIV in both national (up to an 85.75% increase) and local (up to an 83.20% increase) EHR databases than the previously published algorithms. The demographic characteristics of people with HIV identified using our algorithm were similar to those reported in national and local HIV surveillance data. Our algorithm demonstrated improved sensitivity over existing algorithms (98% vs 56%-92%) while maintaining a similar overall accuracy (96% vs 80%-96%).

Conclusions: We developed and evaluated an updated criteria-based phenotyping algorithm for identifying people with HIV in EHR data that demonstrates improved sensitivity over existing algorithms.

(*JMIR Form Res* 2021;5(11):e28620) doi:[10.2196/28620](https://doi.org/10.2196/28620)

KEYWORDS

phenotyping; algorithms; electronic health records; people with HIV; cohort identification

Introduction

Background

The widespread adoption of electronic health records (EHRs) by health care systems over the last decade has led to an explosion in available clinical data. These databases allow researchers to retrospectively study large cohorts of patients with a specific disease or set of clinical characteristics of interest for quality improvement projects and clinical research. However, the increase in the amount of available data brings with it the need for more efficient methods of identifying patient cohorts to facilitate this research. Historically, cohorts were identified via manual chart review, a process by which a trained abstractor manually reviewed each patient record to determine their eligibility for inclusion in the study. However, this is a time- and resource-intensive process and is impractical for large EHR databases containing thousands or millions of patients.

The limitations of manual chart review have led to the emergence of phenotyping algorithms to automate the identification of patient cohorts from large data sets for a variety of conditions such as diabetes, heart disease, and asthma [1-4]. Here, we focus on automated algorithms for identifying cohorts of people with HIV in EHR databases, as such cohorts can be useful for studying engagement along every step of the HIV care continuum (diagnosis, linkage to care, retention in care, and viral suppression) [5] and identifying areas for improvement, including strategies for prevention in high-risk populations.

The earliest algorithms for identifying people with HIV used administrative data from government databases such as those comprising Medicare or Medicaid claims [6-11]. As these data sets generally contain only diagnostic (International Classification of Disease [ICD]) and procedure (Current Procedural Terminology) codes, the cohort definition algorithms for HIV developed for them rely solely on ICD codes. An example of this type of algorithm requires at least 2 outpatient ICD codes for HIV or 1 inpatient ICD code for HIV to classify a patient as having HIV [11]. The reliability of these algorithms is however limited when applied to EHR data where relying only on ICD codes can lead to misclassification of people with HIV if these codes are used incorrectly, for example, using HIV-specific codes for testing or prevention counseling. ICD codes could also be missing as would be the case if the primary reason for the encounter was not for management of HIV infection. Recent studies have sought to improve the performance of ICD code-based algorithms on EHR data by developing phenotyping algorithms that mirror the testing and diagnostic guidelines from the US Centers for Disease Control and Prevention (CDC) [12]. These algorithms use data such as laboratory test results and prescriptions for HIV-specific medications, as well as ICD codes, to identify people with HIV from EHR records [13-18], and demonstrate good sensitivity and specificity. However, they were developed using data from single health care systems or the Department of Veterans' Affairs, which could limit their generalizability.

As stated previously, recent HIV phenotyping algorithms developed for EHR data are based on clinical steps taken to diagnose HIV infection to provide a step-by-step procedure for

identifying people with HIV in the data; for example, first identify all patients with a positive HIV screening test or an ICD code for HIV, then from this set identify those with a positive confirmatory test, etc. Although these algorithms make use of the clinical information contained in the EHR to confirm HIV diagnosis, they can miss people with HIV who do not have complete documentation of their diagnostic history or do not have ICD codes for HIV documented in their records. We found this to be the case when we implemented 2 recently published HIV phenotyping algorithms that follow this model [18] in our data and identified a significant number of people with HIV who had clinical evidence of HIV infection but were missed by these algorithms.

An alternative method that potentially avoids misclassification because of missing data is to develop a set of criteria to define HIV diagnosis. Kramer et al [16] described a set of 3 criteria: (1) presence of an ICD, ninth revision (ICD-9) code for HIV; (2) a positive HIV laboratory test, defined as a positive screening test, positive Western blot, HIV viral load (VL) measurement regardless of result, or CD4 count measurement regardless of result; and (3) prescription for HIV-specific antiretroviral medications at any time. They found that requiring at least 2 of the 3 criteria to classify a patient as having HIV yielded the highest sensitivity, with minimal trade-off in positive predictive value. However, as the algorithm by Kramer et al [16] requires evidence of a VL, without relying on the values of the VL, changes in the guidelines for HIV testing and diagnosis could introduce false positives as VL measurement is increasingly used for diagnostic purposes (especially in the diagnosis of acute HIV infection) and not solely for monitoring infection and treatment [19].

Objective

The objective of this study is to develop and validate a novel phenotyping algorithm to identify people with HIV in EHR data that is based on an updated set of clinical criteria and to capture people with HIV missed by existing algorithms.

Methods

Overview

We implemented and evaluated our new algorithm alongside multiple baseline algorithms for comparison in both a national, multi-institutional EHR database and a local EHR database from a single health care system. Both databases contain data collected before and after the transition from the ICD-9 to the ICD-10, as well as before and after the introduction of new HIV testing guidelines by the CDC. To evaluate the performance of the proposed algorithm, we used 2 different strategies. First, the distribution of several demographic characteristics, such as gender and race or ethnicity, is different among people with HIV than the general patient population [20]. Therefore, we compared the distribution of demographic factors of the cohorts of people with HIV who were identified with those reported by the local health department and the CDC to confirm that our algorithm identified a cohort with representative demographic characteristics. Second, we validated our algorithm in the local EHR database by performing chart review on a subsample of patients (both people with HIV and people without HIV) and

calculating the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy.

Data

Our algorithm was developed and evaluated using both local and national data sets. Our national data set is Cerner HealthFacts, a deidentified EHR database containing records of 69 million unique patients from over 600 participating hospitals and clinics spanning 19 years from 1999 to 2017. Local data were derived from University of Texas (UT) Physicians, an outpatient network based in Houston, Texas. This database contains records of approximately 4 million patients between 2006 and 2020. Both data sets have undergone harmonization and normalization procedures by the Cerner organization (national data) or the UTHealth clinical data warehouse team (local data) to ensure data validity. Patient demographics (gender, race or ethnicity, marital status, and insurance), census region of clinic (for the national database), urban or rural status, diagnosis codes, results of laboratory studies, and medications were extracted from both databases for all patients aged ≥ 13 years. Furthermore, 13 was selected as the age threshold for inclusion in the study because testing guidelines from the CDC recommend beginning screening for HIV at the age of 13 years [19]. The use of these data in this study was approved by the UTHealth Committee for the Protection of Human Subjects.

Baseline and HIV-Phen Phenotyping Algorithms

We implemented 4 previously published HIV phenotyping algorithms in both data sets and used them as baseline comparators for our new algorithm. The first baseline algorithm is based only on ICD codes for HIV described by Fultz et al [11]. This algorithm requires at least 2 ICD codes for HIV documented in an outpatient setting or 1 ICD code for HIV documented in an inpatient setting to classify a patient as a person with HIV. A complete list of ICD-9 and ICD-10 codes required to implement this algorithm can be found in [Multimedia Appendix 1](#) (Table S1).

A total of 2 HIV phenotyping algorithms described by Paul et al [18] were used as the second and third baselines. The second baseline algorithm closely follows CDC testing and diagnostic guidelines and relies on laboratory results and medications to identify people with HIV. This algorithm first identified all patients with a positive HIV antibody screening test ([Multimedia Appendix 1](#); Table S2) and then identified those in this group with a positive HIV confirmatory test (Western blot, immunofluorescence assay, or HIV-1/2 differentiation assay [MultiSpot or Geenius]; [Multimedia Appendix 2](#) [12]) as having a confirmed HIV diagnosis. Patients who did not have a record of an HIV screening test, had a negative or indeterminate screening test, or had a positive screening test and a negative or indeterminate confirmatory test were considered to have HIV infection if they had an HIV VL > 1000 copies/mL. For patients with a VL ≤ 1000 copies/mL or an undetectable VL, their medication history was reviewed for prescriptions for antiretroviral medications for HIV treatment, which would

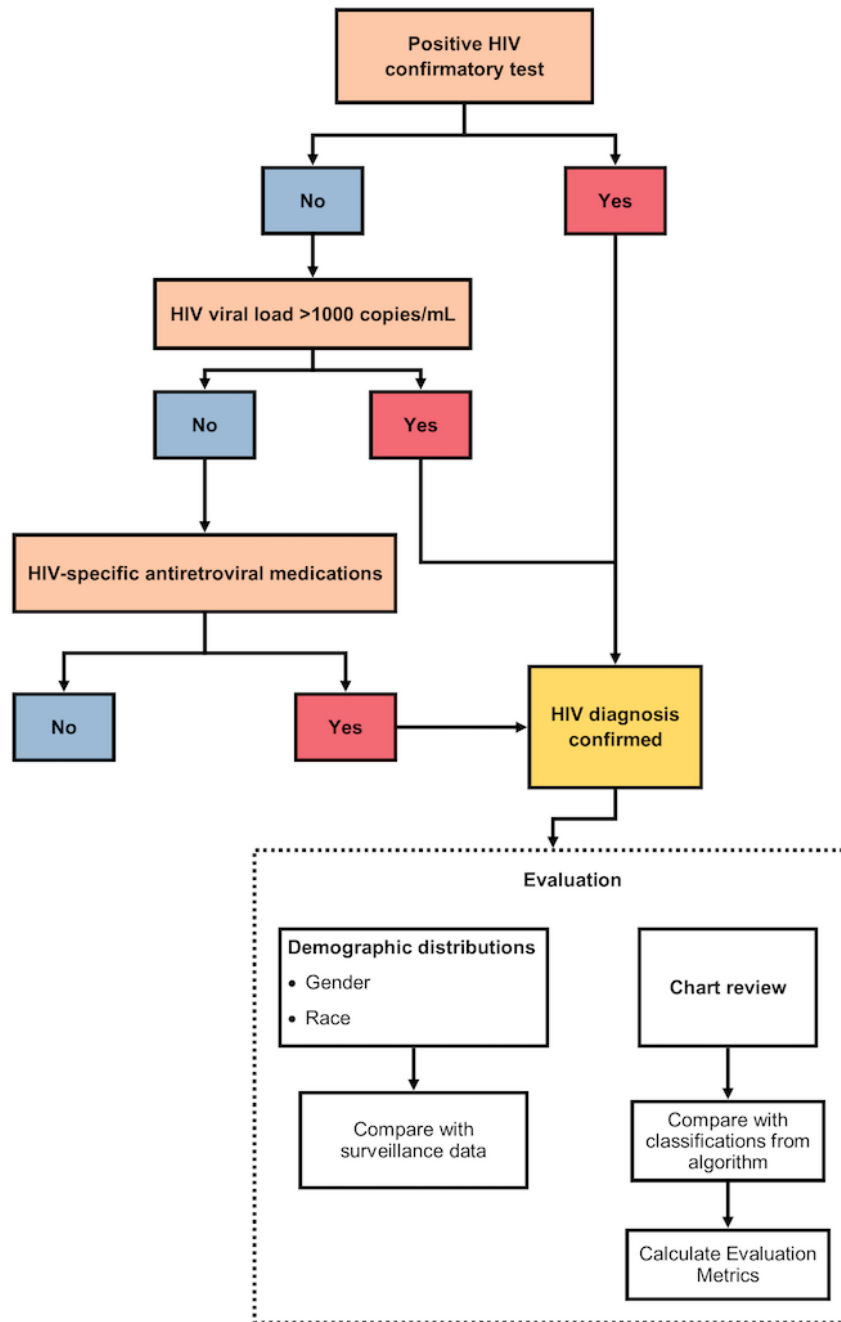
confirm their HIV diagnosis. The lists of HIV VL test and HIV antiretroviral medications used to implement this algorithm can be seen [Multimedia Appendix 2](#).

The third baseline algorithm is ICD code-based and begins by identifying all patients with an ICD-9 or ICD-10 code for HIV or HIV-related comorbidities ([Multimedia Appendix 1](#); Table S1). Patients in this group with a positive HIV confirmatory test ([Multimedia Appendix 2](#)) or HIV VL > 1000 copies/mL ([Multimedia Appendix 2](#)) were considered to have a confirmed HIV diagnosis. Those who did not meet these criteria were reviewed for prescriptions for HIV antiretroviral medications ([Multimedia Appendix 2](#)) to confirm HIV diagnosis.

The fourth and final baseline we implemented as a comparator for our new algorithm was the criteria-based algorithm described by Kramer et al [16]. This algorithm defines a set of 3 criteria and requires that at least 2 of the 3 criteria be met to classify a patient as having HIV. These criteria are (1) presence of an ICD-9 code for HIV, (2) a positive HIV laboratory test, defined as a positive screening test, positive confirmatory test, HIV VL measurement regardless of result, or CD4 count measurement regardless of result, and (3) prescription for HIV-specific antiretroviral medications at any time. The ICD codes needed to implement this algorithm are listed in Table S1 of [Multimedia Appendix 1](#), HIV screening tests are listed in Table S2 of [Multimedia Appendix 1](#), CD4 count tests are listed in Table S3 of [Multimedia Appendix 1](#), HIV confirmatory tests are listed in [Multimedia Appendix 2](#), and HIV VL test are listed [Multimedia Appendix 2](#).

Our new algorithm identifies a minimum set of clinical criteria, only one of which must be met to confirm HIV diagnosis. These criteria are a positive HIV confirmatory test, an HIV VL > 1000 copies/mL, or a prescription for HIV antiretroviral medications sufficient to treat (rather than prevent) HIV as evidence of a confirmed HIV diagnosis. A decision tree representing our phenotyping algorithm is depicted in [Figure 1](#), and the pseudocode that details the data points needed to implement this algorithm can be seen in [Multimedia Appendix 2](#), as well as on Phenotype Knowledgebase (PheKB) [21]. Initial lists of HIV laboratory tests were generated from both the national and local databases using HIV-related keywords (*HIV*, *human immunodeficiency virus*, *rapid*, *multispot*, and *geenius*). The lists had to be generated separately for each database as laboratory test names were not standardized across institutions, and different health care systems often use different names and terminology for the tests. These lists were reviewed by a clinical domain expert (TPG) to generate the final lists of relevant laboratory tests for each data set to be included in the algorithm. In addition, a list of HIV antiretroviral medications used to treat HIV was compiled with the assistance of the same clinical domain expert. Patients being treated with only a subset of HIV antiretroviral medications that can be used to treat hepatitis B infection or for pre-exposure prophylaxis were required to have a positive confirmatory test or a VL > 1000 copies/mL to be considered to have a confirmed HIV diagnosis.

Figure 1. Diagram of our HIV phenotyping algorithm and evaluation framework.



Evaluation

A total of 2 strategies were used to evaluate the performance of our algorithm: a comparison of demographic statistics to national and local statistics and a chart review to validate HIV status. First, comparisons of the demographic distributions between the cohorts of people with HIV identified using our algorithm and those with HIV included in local and national surveillance data were performed to provide evidence that our

algorithm correctly identifies people with HIV from the EHR data. Demographics of the national cohort were compared with national demographic distributions of people with HIV from the HIV Surveillance Report published each year by the CDC [20], and the demographics of the local cohort were compared with demographic distributions from people with HIV in the Houston area compiled by the Houston Area Ryan White Planning Council and Houston Health Department [22].

Second, a random subsample of cases (people with HIV) and control patients was extracted from the UT Physicians data. As HIV cases are infrequent in the data, to maintain a 1:1 ratio of cases to controls, we randomly sampled patients determined as cases or controls by our algorithm. The sample size was calculated based on a 95% confidence level and a margin of error of 5%. Chart review, guided by the clinical domain expert (TPG), was then performed on this subsample to determine the HIV status of each patient by one of the researchers (SBM). On the basis of this chart review–based gold standard, the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were calculated for our algorithm. The evaluation framework is shown in [Figure 1](#). For comparison, these metrics were also calculated for the baseline phenotyping algorithms implemented in the local EHR.

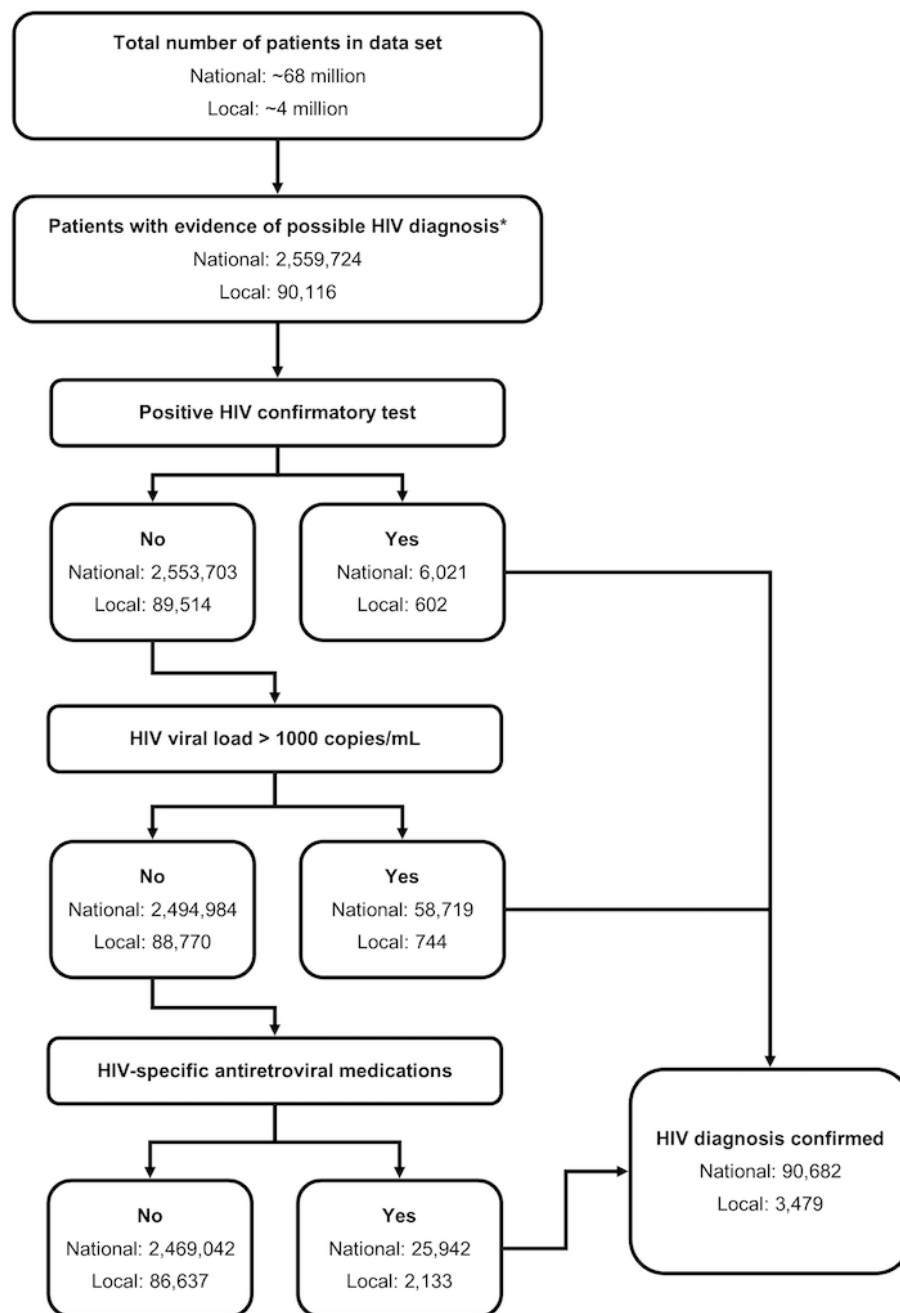
Results

Characteristics of People With HIV Cohorts Identified by the Algorithms

In the national EHR data, the ICD-only baseline identified 86,066 people with HIV, the laboratory-based baseline algorithm identified 65,629 people with HIV, the ICD-based baseline algorithm identified 48,819 people with HIV, and the criteria-based baseline identified 72,443 people with HIV. In contrast, our algorithm identified 90,682 people with HIV. This represents a 5.36%, 38.17%, 85.75%, and 25.18% increase in the number of people with HIV identified in this data set over the baseline algorithms, respectively. A diagram showing the flow of patients using our algorithm is displayed in [Figure 2](#).

We examined how patients qualified as having HIV using our algorithm to identify why it resulted in the identification of more people with HIV. A Venn diagram showing the number of people with HIV identified by each criterion of the new algorithm in the national data set is shown in [Figure 3A](#). Most of the patients in the cohort identified by our algorithm (58,719/90,682, 64.75%) were detected based on the presence of an HIV VL test result >1000 copies/mL, 48.47% (28,463/58,719) of whom did not have an ICD code for HIV or a positive HIV screening test. An additional 6021 patients were included in the cohort based on the presence of a positive HIV confirmatory test, making up 6.64% (6021/90,682) of the national cohort of people with HIV. Of these 6021 patients, 1732 (28.77%) did not have an ICD code for HIV or a positive screening test in the data, leading them to be missed by one or more of the baseline algorithms. Finally, 28.61% (25,942/90,682) were detected based on the presence of a prescription for HIV antiretroviral medications, and 35.75% (9274/25,942) of these patients did not have an ICD code for HIV or a positive screening test for HIV documented in the data. All people with HIV identified by the laboratory-based and ICD-based baseline algorithms were also identified by our new algorithm. However, there were 42,382 patients identified as people with HIV by the ICD-only baseline that were not identified by our algorithm. Conversely, our algorithm identified 46,994 people with HIV that the ICD-only baseline had not identified. The criteria-based baseline identified 22,536 patients as people with HIV not identified by our algorithm, whereas our algorithm identified 40,775 people with HIV not identified by the criteria-based baseline.

Figure 2. Flow diagram of patients through our algorithm for both national and local data sets. *Any International Classification of Disease code for HIV, HIV-related laboratory test performed regardless of result, or medication used to treat HIV documented in the data.



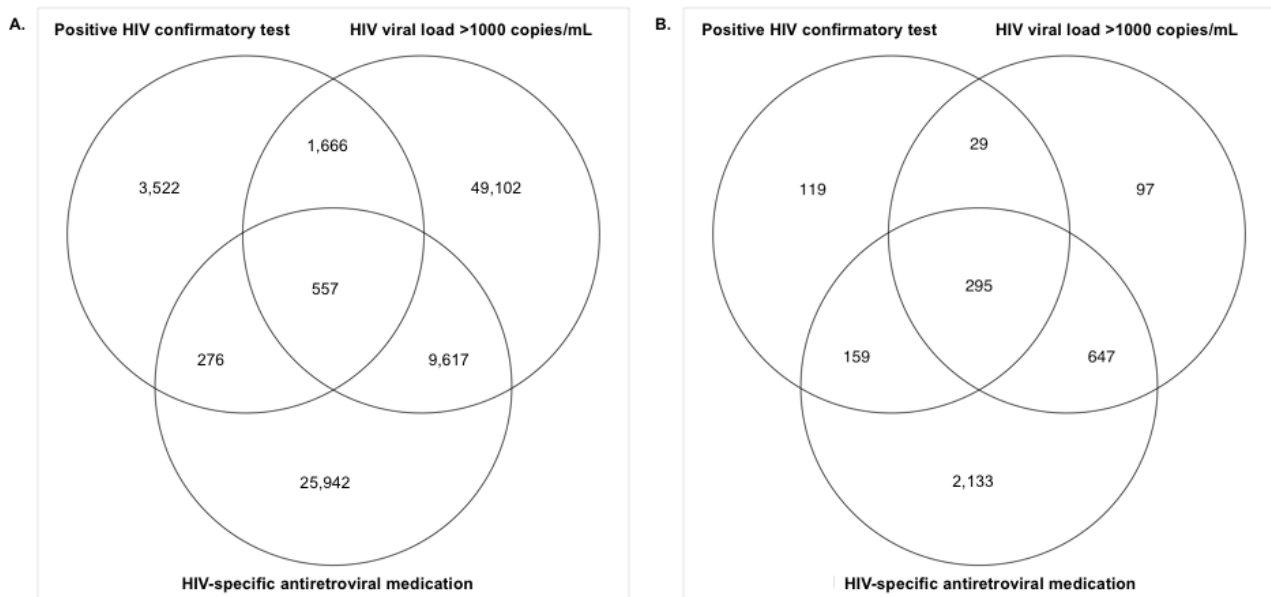
Similar results were obtained from the local database. The baseline algorithms identified 3399, 1899, 2764, and 2911 people with HIV in the local data (ICD-only, laboratory-based, ICD-based, and criteria-based baselines, respectively). This is in comparison with the 3479 people with HIV identified by our algorithm in these data (Figure 2). This represents a 2.35%, 83.20%, 25.87%, and 19.51% increase in the number of people with HIV identified by our new algorithm over the ICD-only,

laboratory-based, ICD-based, and criteria-based baseline algorithms, respectively. Similar to the national cohort, all people with HIV identified by the laboratory-based and ICD-based baseline algorithms were also identified by our algorithm, whereas the ICD-only baseline identified 752 people with HIV not identified by our algorithm and the criteria-based algorithm identified 84 people with HIV not identified by our algorithm. Conversely, our algorithm identified 832 people with

HIV not identified by the ICD-only baseline and 652 people with HIV not identified by the criteria-based baseline. Contrary to the national cohort, most of the local cohort (2133/3479, 61.31%) were identified by the presence of a prescription for HIV antiretroviral medications, 26.91% (574/2133) of whom did not have an ICD code for HIV or a positive HIV screening test in the data. Only 21.39% (744/3479 patients) of the local cohort were identified based on the presence of an HIV VL

result >1000 copies/mL (50/744, 6.7% of whom did not have an ICD code for HIV or a positive HIV screening test), and 17.3% (602/3479 patients) based on a positive confirmatory test (18/602, 3%) of whom did not have an ICD code for HIV or a positive HIV screening test in the data). A Venn diagram showing the number of people with HIV identified by each criterion of the new algorithm in the local data set is shown in Figure 3B.

Figure 3. Venn diagram showing the number of patients meeting each of the criteria of our HIV phenotyping algorithm for (A) national data set, and (B) local data set.

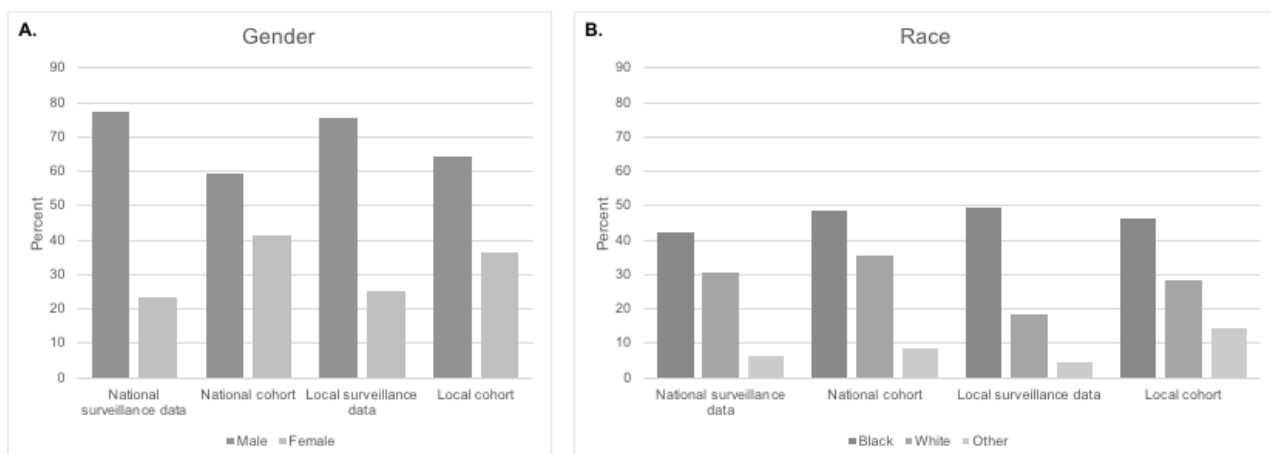


Demographic Characteristics of People With HIV Cohorts Identified Using Our Algorithm

To validate the ability of our algorithm to identify people with HIV in both data sets, we analyzed the distributions of several demographic characteristics and compared these distributions with those seen in local and national HIV surveillance data. The cohort from our national data shares race distributions similar to those reported in national surveillance data of people with HIV: 48.43% (43,915/90,682) of people with HIV are Black

and 34.69% (31,462/90,682) are White in our national data compared with 40.61% (423,304/1,042,270) Black and 29.19% (304,206/1,042,270) White in the national surveillance data (Figure 4B). Our national cohort demonstrated a higher proportion of males than females, which corresponds to the distributions seen in national surveillance data collected by the CDC [20]. However, the percentage of females in our national cohort was much higher (36,738/90,682, 40.51%; Figure 4A) than that reported nationally (245,727/1,042,270, 23.58%; Figure 4A), with a correspondingly lower percentage of males.

Figure 4. Comparison of distributions of gender and race in cohorts identified by our algorithm and national (Centers for Disease Control and Prevention) and local (Houston Health Dept) HIV surveillance data.



In the local EHR data, demographic distributions of people with HIV identified by our algorithm were compared with local surveillance statistics reported by the Houston Health Department for people with HIV in the area [22]. As with the national cohort, the racial distribution of the cohort from our local data (1589/3479, 45.67% Black; 984/3479, 28.28% White) was comparable with the racial distribution reported in local surveillance data (12,424/25,132, 49.43% Black; 4608/25,132, 18.34% White; Figure 4B). Similar to the national cohort, the local cohort was slightly more female than reported in the surveillance data (1239/3479, 35.61% females in our cohort vs 6171/25,132, 24.55% in the surveillance data; Figure 4A).

Evaluation of Our Algorithm Compared With Baseline Algorithms in Local EHR Data

A chart review was performed on a random subsample of 360 patients in the local data set to evaluate the performance of our

HIV phenotyping algorithm compared with the baseline algorithms. The sensitivity of our algorithm was 98%, representing a substantial increase in sensitivity over the laboratory-based baseline algorithm (56%) and ICD-only baseline algorithm (86%), as well as a moderate increase over the ICD-based baseline algorithm (90%) and criteria-based baseline algorithm (92%). In addition, our algorithm demonstrated an increase in overall accuracy over 3 of the 4 baselines (HIV-Phen, 96%; ICD-based baseline, 95%; laboratory-based baseline, 80%; ICD-based baseline, 95%). However, these gains were accompanied by a decrease in the specificity of our algorithm compared with the baseline algorithms. Side-by-side comparisons of these results are presented in Table 1.

Table 1. Evaluation results^a.

Algorithm	Source	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Accuracy	Results			
							C+		C-	
							A+	A-	A+	A-
ICD ^b -only baseline	Fultz et al [11]	0.86	0.99	0.99	0.89	0.93	143	23	1	193
Laboratory-based baseline	Paul et al [18]	0.56	1.0 ^c	1.0	0.73	0.80	93	73	0	194
ICD-based baseline	Paul et al [18]	0.90	1.0	1.0	0.92	0.95	149	17	0	194
Criteria-based baseline	Kramer et al [16]	0.92	0.99	0.99	0.93	0.96	152	14	1	193
HIV-Phen	N/A ^d	0.98	0.95	0.95	0.98	0.96	162	4	9	185

^aResults of the evaluation of the baseline algorithms and our new clinical criteria-based algorithm on a subsample of 360 patients from the local database. Classification of the patients is shown on the right side of the table (Results) broken down by the results of chart review (C+ or C-) and algorithm classification (A+ or A-).

^bICD: International Classification of Disease.

^cResults are italicized for the algorithm with the highest value for each metric.

^dN/A: not applicable.

Discussion

Principal Findings

We developed a novel HIV phenotyping algorithm, HIV-Phen, that relies solely on laboratory and medication data and requires only 1 of 3 clinical criteria to be met to identify people with HIV: positive HIV confirmatory test, HIV VL>1000 copies/mL, or prescription of HIV antiretrovirals for the treatment of HIV. This algorithm was developed to address a significant portion of people with HIV that were missed by previously published algorithms in both our local and national data sets. Our new algorithm is able to identify up to 85.75% more people with HIV in our data and demonstrates improved sensitivity over the baseline comparators with modest trade-off in specificity.

We found that these people with HIV were missed by the baseline algorithms because a number of people with HIV had

laboratory or prescription evidence to confirm HIV diagnosis but did not have ICD codes for HIV documented in their medical records. Furthermore, patients often visit providers from multiple health care systems for care. Owing to this, in a given EHR, some people with HIV might not have all the laboratory information needed to confirm HIV diagnosis according to algorithms that mimic clinical testing and diagnostic guidelines such as the laboratory-based baseline algorithm implemented here. This leads such algorithms to misclassify people with HIV. Our algorithm was implemented and compared across both national and local EHR databases containing data spanning large time scales, including both ICD-9-CM and ICD-10 billing codes, as well as various HIV testing technologies and guidelines. In addition to differences in geographic coverage, the 2 data sets used in this study are very different in composition with the local data containing mostly outpatient data and the national data containing predominantly inpatient as well as outpatient data. The fact that our new algorithm

performs well in both data sets supports its portability across EHRs from different sources.

Our algorithm demonstrates a marked improvement in sensitivity over the baseline algorithms and a small increase in accuracy in the local data. However, these gains come with small decreases in specificity. Our algorithm resulted in 9 false positives, or people who were identified as people with HIV by the algorithm when they did not have HIV. Most patients in this group were falsely identified as having HIV because they were prescribed postexposure prophylaxis (PEP) because of HIV exposure. Unlike patients on pre-exposure prophylaxis or hepatitis B virus treatment, these patients are more difficult to identify and exclude because PEP consists of a full HIV treatment regimen. As PEP is only taken for a single short period (typically 30 days) after exposure, prescription duration and count could be considered in the algorithm to correctly identify these individuals. However, this could exclude people with HIV whose infection is not being managed by a provider in the system, who are nonadherent, or have fallen out of care. Another source of false positives was variations in clinical practices, that is, patients who received prescriptions for antiretrovirals on the same day as a screening test that came back negative.

Our algorithm falsely classified 4 patients as negative when they had HIV infection. This was largely because of HIV infection only being mentioned in the text of clinical notes for these individuals and no laboratories or medications for HIV listed in their records. Owing to this, they were also misclassified by the laboratory-based baseline algorithm. In addition, these patients did not have ICD codes in the EHR and were thus, falsely classified as negative by the other baseline algorithms that make use of these codes. Out of the 4 false negatives, 1 had a detectable HIV VL, but it was <1000 copies/mL, which is not considered sufficient clinical evidence to confirm HIV diagnosis by our criteria. We chose 1000 copies/mL as the cutoff for confirming HIV to accommodate the changes over the years in the sensitivity of the VL test, as over the time frame of our data, the lower limit of detection of the VL test has gone from 400 copies/mL to 48 copies/mL to 20 copies/mL. Running our algorithm in the local data using VL thresholds of 400, 48, and 20 copies/mL increases the number of people with HIV identified by less than 1% over the 1000 copies/mL threshold, and only 1 additional patient was identified as positive in the evaluation subsample. This was the patient with a detectable VL <1000 copies/mL that was falsely classified as negative previously. Running the same analysis in the national data increases the number of people with HIV identified by 1.5%, 7%, and 10% with VL thresholds of 400, 48, and 20 copies/mL, respectively. However, a threshold of >1000 copies/mL reduces the possibility of false positivity in distinguishing acute HIV infection from a false positive screening test when the differentiation assay is negative.

As further evidence of the accuracy of our new algorithm, we found good agreement in most demographic trends with HIV surveillance data in the cohorts from both the local and national data. However, we observed a higher percentage of women in both the national and local cohorts than that reported in the surveillance data. In the local data set, this was observed in cohorts resulting from our algorithm and the baseline algorithms.

Given this, and as this is a cohort derived from clinical data rather than surveillance data, the higher percentage of women could be because of differences in the characteristics of patients who use the UT Physicians network for health care compared with the general population of people with HIV in the Houston area. As most people with HIV in the local cohort were identified because they have a prescription for antiretrovirals but lack HIV laboratory data, we speculate that many of the people with HIV identified are accessing specialty care in the UT Physicians system for other conditions and their HIV infection is being managed elsewhere. Furthermore, studies have shown that women are more likely to consult a physician than men and are more likely to have health insurance and a regular source for health care [23-26], which could potentially drive the discrepancy in gender distribution between the UT Physician population and the Houston-area surveillance data.

A discrepancy in gender distribution was also observed between the national cohort and the national HIV surveillance data. On examining the distribution of gender by census region in our national cohort, we found that the gender distribution of people with HIV from the West and Midwest align very well with the national surveillance data distribution; however, people with HIV from the South and Northeast have a much higher percentage of women than reported in the national surveillance data. Most people with HIV identified by our algorithm resided in these regions. Demographic distributions among people with HIV are not uniform across the country, and regional variations exist; for example, heterosexual transmission is a more predominant risk factor in the South, which results in a higher percentage of people with HIV who are women in this region [20]. These regional variations and the characteristics of people who regularly interact with the health care system mentioned previously could be partly responsible for the differences we observed in gender distribution in the national cohort compared with the national surveillance data.

In both national and local data, the number of Hispanic patients was not accurately captured. This is likely because of differences in the way this information is collected across systems, that is, as a single race or ethnicity variable or as separate race and ethnicity variables. In our data, this likely leads to the number of White patients being overestimated in the data and Hispanic patients being underestimated, which may explain the discrepancies observed between our national and local people with HIV cohorts and national and local surveillance data.

Limitations

Our study has several limitations. First, although the national EHR data set is very large and contains records of millions of patients from across the country, it is deidentified, and thus, lacks information such as clinical notes and identifiers that could be linked to other data sources or used to validate algorithms against medical record reviews. This information could provide a better understanding of why so many people with HIV in this data lack ICD codes for HIV. Second, the national data are aggregated from multiple clinics across the country and mapped to standard ontologies, such as Logical Observation Identifiers Names and Codes for laboratory tests by Cerner. Errors in the mapping led to ambiguity in the data; for example, tests that

were mapped to HIV screening test but had values with ranges that suggested they were VL measurements. Third, although the national data do provide a nongovernmental national sample of patients, it is limited to clinics that have implemented Cerner EHRs, which could introduce bias. Finally, as laboratory test names are not standardized and different clinics use different names and terminology, lists of HIV laboratories must be generated for each data set on which the algorithm is run, which is a time-consuming process. A consistent mapping to a standardized ontology, such as Logical Observation Identifiers Names and Codes, is needed to fully automate this part of the algorithm.

Conclusions

We have developed and evaluated HIV-Phen, a novel HIV phenotyping algorithm, to identify people with HIV in EHR

data that greatly improves sensitivity over previously published algorithms. In addition, we have shown that a significant proportion of people with HIV in 2 clinical data sets do not have ICD codes for HIV and are thus missed by phenotyping algorithms that rely on this information. This work will positively impact future HIV research as our new algorithm can be applied to both single and multi-institutional data sets to accurately identify complete cohorts of people with HIV to facilitate multiple types of studies. Furthermore, this work seeks to provide a blueprint for the implementation of our algorithm to assist other researchers in the identification of their cohorts. Although the elucidation of algorithms to accurately identify specific cohorts of patients is important, further work is needed to define standard mappings of laboratory tests, medications, and other information to increase the ease and speed with which these algorithms can be implemented across different data sets.

Authors' Contributions

SBM conceived the study, performed the analyses, and drafted the manuscript. TPG provided clinical guidance for the study and provided valuable edits and suggestions for the final manuscript. AG supervised the study, provided guidance on analyses and computational aspects, and made valuable edits and suggestions for the final manuscript. SBM is supported by the United States National Library of Medicine Training Program in Biomedical Informatics and Data Science T15LM007093. TPG is supported by the MD Anderson Foundation Chair at Baylor College of Medicine.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Features required for baseline algorithms.

[DOCX File, 21 KB - [formative_v5i11e28620_app1.docx](#)]

Multimedia Appendix 2

HIV phenotyping algorithm (HIV-Phen) pseudocode.

[DOCX File, 173 KB - [formative_v5i11e28620_app2.docx](#)]

References

1. Afshar M, Press VG, Robison RG, Kho AN, Bandi S, Biswas A, et al. A computable phenotype for asthma case identification in adult and pediatric patients: external validation in the Chicago Area Patient-Outcomes Research Network (CAPriCORN). *J Asthma* 2017 Oct 13:1-8. [doi: [10.1080/02770903.2017.1389952](#)] [Medline: [29027824](#)]
2. Denaxas S, Gonzalez-Izquierdo A, Direk K, Fitzpatrick NK, Fatemifar G, Banerjee A, et al. UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER. *J Am Med Inform Assoc* 2019 Dec 01;26(12):1545-1559 [FREE Full text] [doi: [10.1093/jamia/ocz105](#)] [Medline: [31329239](#)]
3. Pasea L, Chung S, Pujades-Rodriguez M, Shah AD, Alvarez-Madrado S, Allan V, et al. Bleeding in cardiac patients prescribed antithrombotic drugs: electronic health record phenotyping algorithms, incidence, trends and prognosis. *BMC Med* 2019 Nov 20;17(1):206 [FREE Full text] [doi: [10.1186/s12916-019-1438-y](#)] [Medline: [31744503](#)]
4. Spratt SE, Pereira K, Granger B, Batch BC, Phelan M, Pencina M, DDC Phenotype Group, et al. Assessing electronic health record phenotypes against gold-standard diagnostic criteria for diabetes mellitus. *J Am Med Inform Assoc* 2017 Apr 01;24(e1):121-128 [FREE Full text] [doi: [10.1093/jamia/ocw123](#)] [Medline: [27616701](#)]
5. Mugavero MJ, Amico KR, Horn T, Thompson MA. The state of engagement in HIV care in the United States: from cascade to continuum to control. *Clin Infect Dis* 2013 Oct;57(8):1164-1171. [doi: [10.1093/cid/cit420](#)] [Medline: [23797289](#)]
6. Keyes M, Andrews R, Mason ML. A methodology for building an AIDS research file using Medicaid claims and administrative data bases. *J Acquir Immune Defic Syndr* (1988) 1991;4(10):1015-1024. [Medline: [1832459](#)]
7. Peabody JW, Luck J, Jain S, Bertenthal D, Glassman P. Assessing the accuracy of administrative data in health information systems. *Med Care* 2004 Nov;42(11):1066-1072. [doi: [10.1097/00005650-200411000-00005](#)] [Medline: [15586833](#)]
8. Fasciano NJ, Cherlow AL, Turner BJ, Thornton CV. Profile of Medicare beneficiaries with AIDS: application of an AIDS casefinding algorithm. *Health Care Financ Rev* 1998;19(3):19-38. [Medline: [10345411](#)]

9. Thornton C, United S. Methods for identifying AIDS cases in Medicare and Medicaid claims data. Health Care Financing Administration; Mathematica Policy Research, Inc, United States. 1997. URL: <http://archive.org/details/methodsforidenti00thor> [accessed 2020-07-23]
10. Macinski SE, Gunn J, Goyal M, Neighbors C, Yerneni R, Anderson BJ. Validation of an optimized algorithm for identifying persons living with diagnosed HIV from New York State Medicaid Data, 2006-2014. *Am J Epidemiol* 2020 May 05;189(5):470-480 [FREE Full text] [doi: [10.1093/aje/kwz225](https://doi.org/10.1093/aje/kwz225)] [Medline: [31612200](https://pubmed.ncbi.nlm.nih.gov/31612200/)]
11. Fultz SL, Skanderson M, Mole L, Gandhi N, Bryant K, Crystal S, et al. Development and verification of a "virtual" cohort using the National VA Health Information System. *Med Care* 2006 Aug;44(8 Suppl 2):25-30. [doi: [10.1097/01.mlr.0000223670.00890.74](https://doi.org/10.1097/01.mlr.0000223670.00890.74)] [Medline: [16849965](https://pubmed.ncbi.nlm.nih.gov/16849965/)]
12. Branson BM, Owen SM, Wesolowski LG, Bennett B, Werner BG, Wroblewski KE, Centers for Disease Control and Prevention (U.S.), Association of Public Health Laboratories, National Center for HIV/AIDS, Viral Hepatitis, and TB Prevention (U.S.). Division of HIV/AIDS Prevention. Laboratory testing for the diagnosis of HIV infection : updated recommendations. Centers for Disease Control and Prevention. 2014. URL: <https://stacks.cdc.gov/view/cdc/23447> [accessed 2021-11-06]
13. Levison J, Triant V, Losina E, Keefe K, Freedberg K, Regan S. Development and validation of a computer-based algorithm to identify foreign-born patients with HIV infection from the electronic medical record. *Appl Clin Inform* 2017 Dec 21;05(02):557-570. [doi: [10.4338/aci-2014-02-ra-0013](https://doi.org/10.4338/aci-2014-02-ra-0013)]
14. Felsen UR, Bellin EY, Cunningham CO, Zingman BS. Development of an electronic medical record-based algorithm to identify patients with unknown HIV status. *AIDS Care* 2014 Apr 30;26(10):1318-1325 [FREE Full text] [doi: [10.1080/09540121.2014.911813](https://doi.org/10.1080/09540121.2014.911813)] [Medline: [24779521](https://pubmed.ncbi.nlm.nih.gov/24779521/)]
15. McGinnis KA, Fine MJ, Sharma RK, Skanderson M, Wagner JH, Rodriguez-Barradas MC, Veterans Aging Cohort 3-Site Study (VACS 3). Understanding racial disparities in HIV using data from the veterans aging cohort 3-site study and VA administrative data. *Am J Public Health* 2003 Oct;93(10):1728-1733. [doi: [10.2105/ajph.93.10.1728](https://doi.org/10.2105/ajph.93.10.1728)] [Medline: [14534229](https://pubmed.ncbi.nlm.nih.gov/14534229/)]
16. Kramer J, Hartman C, White D, Roysse K, Richardson P, Thrift A, et al. Validation of HIV-infected cohort identification using automated clinical data in the Department of Veterans Affairs. *HIV Med* 2019 Sep 26;20(8):567-570 [FREE Full text] [doi: [10.1111/hiv.12757](https://doi.org/10.1111/hiv.12757)] [Medline: [31131549](https://pubmed.ncbi.nlm.nih.gov/31131549/)]
17. Goetz MB, Hoang T, Kan VL, Rimland D, Rodriguez-Barradas M. Development and validation of an algorithm to identify patients newly diagnosed with HIV infection from electronic health records. *AIDS Res Hum Retroviruses* 2014 Jul;30(7):626-633. [doi: [10.1089/aid.2013.0287](https://doi.org/10.1089/aid.2013.0287)]
18. Paul DW, Neely N, Clement M, Riley I, Al-Hegelan M, Phelan M, et al. Development and validation of an electronic medical record (EMR)-based computed phenotype of HIV-1 infection. *J Am Med Inform Assoc* 2018 Feb 01;25(2):150-157 [FREE Full text] [doi: [10.1093/jamia/ocx061](https://doi.org/10.1093/jamia/ocx061)] [Medline: [28645207](https://pubmed.ncbi.nlm.nih.gov/28645207/)]
19. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5514a1.htm> [accessed 2020-04-05]
20. HIV Surveillance Report, 2018 (Preliminary). Centers for Disease Control and Prevention. 2019. URL: <http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html> [accessed 2020-04-06]
21. May SB, Giordano TP, Gottlieb A. HIV. University of Texas Health Science Center at Houston. 2021. URL: <https://phekb.org/phenotype/1628> [accessed 2021-11-06]
22. The 2019 Houston area integrated epidemiologic profile for HIV prevention and care services planning. Houston Area Ryan White Planning Council and Houston Health Department. 2019. URL: <http://www.houstontx.gov/health/HIV-STD/Documents/2019-Epi-Profile-Final-20191212.pdf> [accessed 2020-03-13]
23. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? An analysis of routinely collected UK general practice data. *BMJ Open* 2013 Aug 19;3(8):e003320 [FREE Full text] [doi: [10.1136/bmjopen-2013-003320](https://doi.org/10.1136/bmjopen-2013-003320)] [Medline: [23959757](https://pubmed.ncbi.nlm.nih.gov/23959757/)]
24. Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services. *J Fam Pract* 2000 Feb;49(2):147-152. [Medline: [10718692](https://pubmed.ncbi.nlm.nih.gov/10718692/)]
25. Thompson AE, Anisimowicz Y, Miedema B, Hogg W, Wodchis WP, Aubrey-Bassler K. The influence of gender and other patient characteristics on health care-seeking behaviour: a QUALICOPC study. *BMC Fam Pract* 2016 Mar 31;17:38 [FREE Full text] [doi: [10.1186/s12875-016-0440-0](https://doi.org/10.1186/s12875-016-0440-0)] [Medline: [27036116](https://pubmed.ncbi.nlm.nih.gov/27036116/)]
26. Women's health insurance coverage. Kaiser Family Foundation. 2020. URL: <https://www.kff.org/womens-health-policy/fact-sheet/womens-health-insurance-coverage-fact-sheet/> [accessed 2020-10-12]

Abbreviations

- CDC:** Centers for Disease Control and Prevention
- EHR:** electronic health record
- ICD:** International Classification of Disease
- PEP:** postexposure prophylaxis
- UT:** University of Texas

VL: viral load

Edited by G Eysenbach; submitted 08.03.21; peer-reviewed by V Curcin, Y Chu; comments to author 21.06.21; revised version received 10.08.21; accepted 07.10.21; published 25.11.21.

Please cite as:

May SB, Giordano TP, Gottlieb A

A Phenotyping Algorithm to Identify People With HIV in Electronic Health Record Data (HIV-Phen): Development and Evaluation Study

JMIR Form Res 2021;5(11):e28620

URL: <https://formative.jmir.org/2021/11/e28620>

doi: [10.2196/28620](https://doi.org/10.2196/28620)

PMID: [34842532](https://pubmed.ncbi.nlm.nih.gov/34842532/)

©Sarah B May, Thomas P Giordano, Assaf Gottlieb. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.11.2021. 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

mHealth-Based Health Promotion Intervention to Improve Use of Maternity Care Services Among Women in Rural Southwestern Uganda: Iterative Development Study

Esther Cathlyn Atukunda¹, PhD, MPH; Lynn T Matthews², MPH, MD; Angella Musiimenta¹, PhD; Godfrey Rwambuka Mugenyi¹, MD; Samuel Mugisha^{1,3}, MSc; Norma C Ware⁴, PhD; Celestino Obua¹, PhD; Mark J Siedner^{5,6}, MPH, MD

¹Mbarara University of Science and Technology, Faculty of Medicine, Mbarara City, Uganda

²Division of Infectious Diseases, School of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States

³Innovation Streams Limited (iStreams), Mbarara, Uganda

⁴Global Health and Social Medicine, Harvard Medical School, Harvard University, Boston, MA, United States

⁵Department of Medicine and Center for Global Health, Massachusetts General Hospital, Boston, MA, United States

⁶Harvard Medical School, Harvard University, Boston, MA, United States

Corresponding Author:

Esther Cathlyn Atukunda, PhD, MPH

Mbarara University of Science and Technology

Faculty of Medicine

Plot 11-18, Kabale Road

P O Box 1410

Mbarara City

Uganda

Phone: 256 702949832

Email: eatukunda@must.ac.ug

Abstract

Background: Antenatal care (ANC) prevents perinatal morbidity and mortality, but use of these services in Uganda remains low and maternal mortality rates are among the highest in the world. There is growing evidence that mobile health (mHealth) approaches improve timely communication of health-related information and produce positive health behavior change as well as health outcomes. However, there are limited data to guide development of such interventions in settings where ANC attendance and uptake of skilled maternity care are low.

Objective: The aim of this study is to develop a novel patient-centered mHealth intervention to encourage and support women to use maternity care services in Mbarara district, southwestern Uganda.

Methods: Using an iterative development approach, we conducted formative stakeholder interviews with 30 women and 5 health care providers (HCPs) to identify preferred key ANC topics and characterize the preferred messaging intervention; developed content for SMS text messaging and audio messaging with the help of 4 medical experts based on the identified topics; designed an app prototype through partnership with an mHealth development company; and pilot-tested the prototype and sought user experiences and feedback to refine the intervention through 3 sets of iterative interviews, a focus group discussion, and 5 cognitive interviews. Qualitative data were coded and analyzed using NVivo (version 12.0; QSR International).

Results: Of the 75 women who completed interviews during the development of the prototype, 39 (52%) had at least a primary education and 75 (100%) had access to a mobile phone. The formative interviews identified 20 preferred perinatal health topics, ranging from native medicine use to comorbid disorders and danger signs during pregnancy. In all, 6 additional topics were identified by the interviewed HCPs, including birth preparedness, skilled delivery, male partner's involvement, HCP interaction, immunization, and caring for the baby. Positive audio messaging and SMS text messaging content without authoritative tones was developed as characterized by the interviewed women. The postpilot iterative interviews and focus group discussion revealed a preference for customized messaging, reflecting an individual need to be *included* and *connected*. The women preferred short, concise, clear actionable messages that guided, supported, and motivated them to *keep alert* and seek professional help. Complementary weekly reminders to the women's significant others were also preferred to encourage continuity or prompt the needed social support for care seeking.

Conclusions: We used an iterative approach with diffuse stakeholders to develop a patient-centered audio messaging and SMS text messaging app designed to communicate important targeted health-related information and support rural pregnant women in southwestern Uganda. Involving both HCPs and end users in developing and formulating the mHealth intervention allowed us to tailor the intervention characteristics to the women's preferences. Future work will address the feasibility, acceptability, and effectiveness of this design approach.

(*JMIR Form Res* 2021;5(11):e29214) doi:[10.2196/29214](https://doi.org/10.2196/29214)

KEYWORDS

mHealth app; app development; messaging; health education; health promotion; mobile phone

Introduction

Background

Antenatal care (ANC) reduces perinatal and maternal morbidity and mortality by detecting and treating prenatal complications and identifying women classified as high risk to ensure delivery in skilled settings [1-5]. ANC also provides an opportunity to support women, families, and communities at a critical time in the course of a woman's life [3]. However, the use of perinatal services in Uganda remains low, with correspondingly high rates of unskilled home deliveries [6].

To avert maternal and perinatal deaths, the World Health Organization has called for the development and evaluation of adaptable and context-specific health solutions to promote ANC uptake, including interventions that involve delivering health care or medicine practice over mobile devices (mobile health [mHealth]) to empower women to overcome barriers to care [3]. A mechanism by which mHealth apps might affect positive health benefits is through engendering social support for the end user. Previous studies have observed a significant positive relationship among perceived social support, health care seeking, breastfeeding practices, and infant care practices among mothers [7,8]. Social support can mitigate structural and physical barriers to health and facilitate self-efficacy to complete positive health behaviors [9,10]. Several studies have found mobile phone-based messages to be motivational or inspirational or to offer a source of social support [11], cues to action [12], or a source to challenge and debunk negative beliefs [13], leading to the desired change. However, despite the successes in pilot studies elsewhere, enthusiasm from the public sector, and nearly ubiquitous availability of mobile phones in Uganda [14], mHealth interventions to motivate improved outcomes among pregnant women have not been adopted on a larger scale [15].

In all, 2 systematic reviews have identified methodological issues with prior work, including ambiguous descriptions of the interventions and their mechanisms of impact, which have generally neglected to base interventions on behavioral theory [5,16]. Others have alluded to the vital importance of the content of the message in supporting and affecting behavior change [17]. Therefore, mHealth interventions present an opportunity to address barriers to health care use through a multipronged approach by (1) teaching positive health behaviors and addressing specific health concerns (predisposing factors), (2) empowering and strengthening informed decision-making (enabling factors), and (3) improving the perceived need for use of the available services [5,16,18,19]. Others have suggested that mHealth interventions such as SMS text messaging,

preloaded (or preinstalled) apps, and voice- and web-based portals may help individuals to improve eHealth literacy, internalize benefits of health services, and function as a decision-support tool at the point of care [20-22]. These interventions enhance and support healthier lifestyles, empower or enable individuals to seek help, address specific health concerns, change behavior patterns, and strengthen informed decision-making. mHealth interventions may also strengthen social relationships and support positive health behavior through SMS text messaging reminders [5,16,18,19,23].

Objective

Despite the vast research on recommendations for optimized mHealth care, there are insufficient published data on the mHealth app development process. The development of many mHealth apps has been led by developers and investigators, with limited input from end users or with input restricted to postintervention consumer satisfaction ratings such as *like or dislike* to assess usability [24]. A participatory design process that considers clients' and caregivers' needs and expectations from the development phase may improve uptake and sustainability of mHealth interventions. In this study, we describe the development of a novel automated SMS text messaging and audio messaging patient-centered app designed to motivate and support women to present for ANC in rural southwestern Uganda and, ultimately, opt for skilled delivery. Our overarching aim for this paper is to describe an iterative design process usable for other groups that are designing and implementing similar interventions to promote perinatal health in low-literacy, resource-poor settings.

Methods

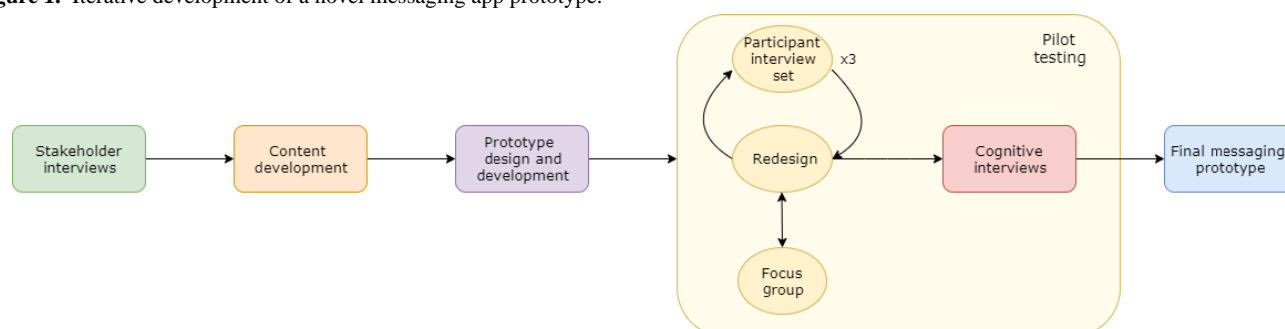
Study Design

We used an iterative design (Figure 1) that comprised the following steps: (1) stakeholder interviews (end-user women and maternal health care providers [HCPs]) to identify key health education topics relevant to the ANC period and characterize women's preferences for an mHealth-based, social support intervention (formative interviews with predevelopment end users); (2) content development for SMS text messaging and audio messaging with the input of 4 medical experts (2 obstetricians and 2 midwives) for the ANC topics as identified and characterized by the stakeholders; (3) design of an app prototype through partnership with an mHealth development company; and (4) pilot-testing the prototype and obtaining feedback for content refinement through (i) 3 sets of iterative exit interviews (pilot participants), (ii) a focus group discussion

(FGD), and (iii) cognitive interviews to further explore user experiences and refine the updated message components and

maximize potential impact and sustained use by rural pregnant women.

Figure 1. Iterative development of a novel messaging app prototype.



Setting

Stakeholder Interviews

We conducted in-depth qualitative interviews with 2 groups of stakeholders: (1) target end users and (2) HCPs. A total of 40 women were invited to complete in-depth interviews, but saturation was attained after 75% (30/40) were interviewed. The women were purposively selected from rural Mbarara district, southwestern Uganda, during the period December 2018 to March 2019. These women were interviewed to identify key health education topics relevant to the ANC period as well as characteristics of a preferred mHealth app. The women were recruited from 10 villages located within 20 km of Mbarara Regional Referral Hospital (MRRH) with the help of existing village health teams (VHTs). VHTs are composed of community-based volunteers who are identified by community members and given basic training on major health programs to mobilize and sensitize communities to use available health services [25-27]. Eligible women for these predevelopment interviews included adults (1) aged ≥ 18 years (2) who had delivered a baby within the past 3 months, (3) owned or had access to a mobile phone, and (4) were able and willing to give informed consent. The purposeful sample was intended to represent women with differing experiences of pregnancy and ANC and included 15 women who delivered at home and 15 who delivered at a health facility.

The interviews were open-ended and organized to cover predesignated core topics. An interview guide was developed and pilot-tested using the constructs of the Healthcare Service Utilization Model as reported elsewhere [26,27] and the Technology Acceptance Model [28,29] (Multimedia Appendix 1). This open-ended approach ensured systematic coverage of specific areas of interest while allowing for unanticipated content to emerge. The interview topics included information and preferred ANC topics considered useful in supporting women during their pregnancy journey, attitude toward using mHealth technology, performance expectancy, effort expectancy, social influence, facilitating conditions, self-efficacy, anxiety, behavioral intention to use technology, and potential technology engagement or fatigue. Specific data on preferences for messaging, content, frequency, preferred language, length, and timing were also sought. A brief questionnaire at the outset of each interview was administered

to collect demographic information (eg, age, occupation, and educational background).

In all, 5 key HCPs (2 obstetricians and 3 midwives) were purposively identified from MRRH and another rural maternity health center in Mbarara district. The HCPs included (1) adults aged ≥ 18 years (2) who were actively engaged in maternity care or policy implementation in Uganda or both, (3) had at least 5 years of experience as HCPs at a busy maternity center, and (4) were able and willing to give informed consent. They were interviewed to explore key ANC health education topics and information to inform the development of the mHealth-based app.

All interviews took place at a private location mutually agreed upon by the participant and the interviewer. Each interview lasted 50-70 minutes. Written informed consent was obtained at the outset of each interview session. Qualitative interviews were digitally recorded with the participant's permission and transcribed verbatim.

Content Development

Using the formative qualitative interviews with the end users, we identified key health education topics along with ANC messages that could be developed to increase *ownership*, engagement, usability, and acceptability by the intended recipients [28,29]. The HCPs identified additional topics that were a critical part of the health education framework during ANC visits [30].

We used Behavior Change Technique Taxonomy version 1 (BCTTv1) [17,31] because it offered a reliable structure to identify, define, interpret, and characterize key components (active ingredients) of our intervention messages aimed at improving the use of maternity care services. Information was grouped within the BCTTv1 components identified as follows: (1) goal setting (outcomes: improving health-related knowledge and skilled delivery), (2) goal setting (behavior: presenting for ANC and avoiding risky behavior), (3) action planning (planning for scheduled visits, financing, actionable messages, partner involvement, and birth preparedness), (4) feedback on behavior (progress monitoring and app interaction features), (5) prompts or cues (follow-up messages or reminders and information cues, eg, danger signs during pregnancy), (6) credible source of information (systematic content development using experts and continuity of care through regular customized information), (7)

instruction on how to perform the behavior, (8) information about health consequences (cautionary social and emotional consequences of, or regrets related to, poor health-seeking behavior), (9) what to do regarding, or where to seek, care or redress (problem solving), (10) review goals (interaction with HCPs and how to review progress), (11) embedded self-monitoring information on progress or preparedness, and (12) active social support for users through regular reminders (self-reminders or through identified social networks).

In all, 4 medical experts (2 obstetricians and 2 midwives), different from those interviewed, were engaged to contribute content for the first draft of the SMS text messages and audio messages, identified by the women and the HCPs, to ensure quality and consistency. The content and frequency of these messages were also based on the type of phone, network, preference, and need to ensure effectiveness and usefulness as well as avoid repetition, fatigue, and burdensomeness.

App Prototype Design and Development

During the intervention design, we worked with *iStreams*, an mHealth app development company in Mbarara town with an existing mHealth platform in Uganda [32]. This local developer designed an initial and novel app prototype that included both SMS text messaging and audio messaging for pilot testing. The design allows training manuals and behavior change communication materials to be integrated within a mobile app. Its unique multimedia design also allows women to listen to messages in their own language or view culturally relevant visuals (such as those that identify danger signs or complications, getting prepared for delivery, expected date of delivery, childbirth checklist, and others). Women were able to register on this platform and be tracked throughout their

pregnancy and postpartum periods, receive automatic or scheduled SMS text message reminders, SMS text messages, audio messages, or notifications about upcoming appointments. In addition, the app stores medical information and allows real-time submission of data directly from a mobile phone, allowing managers and supervisors to access up-to-date data on health outcomes. The elements, content, and patterns of the SMS text messaging reminders were customized and the prototype presented as an *eBirth* platform for SMS text messaging and audio messaging (Figure 2). All messages were developed in English and then translated into the local language, Runyankole, by an experienced translator to ensure that context was maintained. Messages were dispensed in either English or Runyankole as preferred by the recipient. Fixed SMS text messaging data were stored in a secure cloud with *iStreams*, which is Health Insurance Portability and Accountability Act compliant.

We used the *Bendixen* approach [33] for designing and developing a user-centered mHealth app, considering 5 overarching goals: (1) ease of use, (2) engagement, (3) education and preparation, (4) motivation and support, and (5) tailoring the system and personalizing the information for end users. Messages were intended to communicate information on the benefits of nutrition, exercise, presenting for ANC, skilled delivery, partner involvement, birth preparedness, monitoring danger signs, and overcoming barriers to access maternity service. The topics of these messages were identified and characterized by both the women (end users) and the HCPs during the predevelopment stage, and the content was developed by health experts. Scheduled SMS text messaging reminders were incorporated as part of the intervention as a stimulus, prompt, or cue to take action.

Figure 2. The e-Birth app biodata registration form interface.

The image shows a screenshot of the 'Bio-data Registration Form' interface. The form is titled 'Bio-data Registration Form' and is part of the 'Participant Registration' section. It contains various input fields for personal and contact information, including Study ID, First Name, Age, Address, Home District, Relative Contact 1 and 2, Preferred Time Of Call, Gender, Call Initiation Day(s), Communication Choice, Phone Number, Start Date Of Call, Preferred Language, Passcode, Active status, and Close Date. The form is framed by a light gray border with navigation tabs at the top: 'Attribute Generation', 'Participant Registration' (highlighted), and 'Participant Records'. At the bottom, there are 'Submit' and 'Cancel' buttons.

Prototype Pilot Testing

Through the VHTs, we screened and enrolled 30 pregnant women (3 successive iterations of 10 pregnant women) from communities residing within 20 km of MRRH who had not presented for ANC by the beginning of their third trimester (determined by their last menstrual period) to test and assess preliminary feasibility, acceptability, and usability of the novel app through postuse qualitative interviews. An iterative approach of interviewing 10 women in each of the 3 groups was also considered sufficient to obtain rich, specific, and purposefully focused information from participants who had been exposed to the intervention [34]. These novel messages and their content were also tested to ensure ownership, relevance, consistency, and expectations among these pregnant women.

After enrollment, messages were sent to these 30 women (audio messages or SMS text messages or both), depending on the participant's choice. Women with access to a mobile phone in their household were registered, and they received the current version of the messages through the *eBirth* app prototype (Figure 2) for at least 3 months, a period of time chosen to include a minimum of 3 ANC visits and delivery. Messages were sent in a specific sequence, depending on the month of pregnancy, to cover appropriate topics identified in the formative interviews. In the case of some women, SMS text messaging reminders were incorporated and sent to their significant others or social supporters to help remind them and support them on their upcoming ANC visit and maternity journey [23]. However, the social supporters were not given any prior recommendations or instructions guiding them on how to respond to the SMS text messaging reminders because the intervention was designed to

provide social support by building on already existing supportive relationships of the study participants.

To test for frequency preferences, we first sent out messages daily for 2 weeks, then weekly at the chosen times, and then twice a week alternating between 8 AM and 8 PM. A message delivery log was monitored on the app platform. We followed the enrolled women through delivery. Upon completion of the 3-month message delivery period, we interviewed the women using semistructured questionnaires to obtain feedback on content, preferred terminologies, language, and ease of use in obtaining the needed support. We assessed phone use and responsiveness by how often the women read the SMS text messages, received calls, texted back or texted at all, confirmation of receipt, phone calls made, and the times when they missed calls or did not read the SMS text messages sent to them. The women were interviewed on technology acceptance, performance and effort expectations, whether they preferred SMS text messaging or audio messaging as a medium of information delivery, their attitude toward SMS text messaging or audio messaging technology, other preferred terminologies, content scheduling, facilitators, technology engagement, convenience, social influence, facilitating factors, anxiety, need for help using the app (self-efficacy), behavioral intention to use, and preliminary feasibility (network challenges, phone ownership, battery life, resources, frequency, and timing).

The app prototype was modified based on feedback from each iterative round of the pilot interviews. After the third modification, of the 30 participants, 10 (33%) who had had similar exposures to the intervention were randomly recruited from the pilot to constitute an FGD aimed at further refining the relevant message components and helping to limit or prioritize the number of topics included in the messaging app as recommended [35]. Finally, 5 cognitive interviews with a new set of women were conducted to further refine the updated messages and maximize potential impact and sustained use by rural pregnant women.

Data Analysis

We described demographic and clinical data for all qualitative, iterative, and FGD interview participants using standard descriptive statistics. Qualitative analysis began with repeated review of the initial transcripts to identify relevant topics of ANC care, as well as characterization of a preferred mHealth app. Qualitative data were coded with the aid of the data management software, NVivo (version 12.0; QSR International). Coded data were iteratively reviewed and sorted to identify repeated themes (topics) arising from the data. Themes were generated using inductive content analysis [36]. The suggested content consisted of descriptive labels that defined and specified each theme (topic) meaning, along with illustrative quotes taken from the qualitative interviews. Themes were harmonized to be

inclusive throughout all the development stages. Coding was guided by questions about attitude, perceived importance, usefulness, responsiveness, preliminary feasibility, experience with the messaging app, and suggested changes. Negative, positive, and neutral perceptions as well as attitudes were also identified and coded. Data analysis was performed jointly by ECA, GRM, and JN. Both JN and ECA double-coded 5 sampled transcripts, yielding a Cohen κ of 0.796. Together with GRM, we resolved disagreements until we were satisfied with the consistency in our coding to generate a codebook. We aimed at ensuring consistency in coding.

For the iterative interviews, data for frequency, timing, and frequency of messaging during the iterative testing were described using Stata software (version 12.0; StataCorp).

Ethics Statement

All personnel involved in the project had relevant training in human subject research ethics. The study was reviewed and approved by the institutional ethics review committees of Mbarara University of Science and Technology and the Uganda National Council for Science and Technology, Kampala, Uganda. All consenting participants gave written informed consent before study enrollment; in the case of those who could not write, a thumbprint was obtained on the consent form as approved by the ethics committees.

Results

Stakeholder Interviews

A total of 30 women participated in the formative qualitative in-depth interviews. In addition, there were 3 iterative groups of 10 women each (30 pilot participants), 1 FGD with 10 women from the pilot, and 5 new individual cognitive interviews. Ultimately, 75 women were separately involved in the development, refining, and testing of the message content for this app. The median age of the women interviewed was 28 (IQR 24-35) years. Of the 75 women who completed interviews, 39 (52%) had at least a primary education and 75 (100%) had access to a mobile phone and previous experience with SMS text messaging or receiving a phone call. Of the 75 women, only 15 (20%) had access to an Android phone. All women were able to receive and initiate an audio call. Of the 75 women, 62 (83%) were able to read and send SMS text messages in English or Runyankole and 49 (65%) preferred both audio messaging and SMS text messaging as a medium of message delivery. Twice weekly (35/75, 47%) and weekly (34/75, 45%) messages were preferred. Of the 75 participants interviewed, 48 (64%) preferred messages sent before 8 AM and 75 (100%) preferred audio calls lasting 1-2 minutes. The rest of the demographic characteristics are presented in [Table 1](#).

Table 1. Demographic characteristics of the study participants (N=75).

Characteristic	Formative interviews (n=30)	Iterative interviews (n=30)	Focus group discussion (n=10)	All participants (n=75) ^a
Age (years), median (IQR)	26 (20-33)	26 (21-34)	27 (21-34)	28 (24-35)
Education level >primary school, n (%)	12 (40)	15 (50)	4 (40)	39 (52)
Access to a mobile phone, n (%)	30 (100)	30 (100)	10 (100)	75 (100)
Experience with SMS text messaging, n (%)	30 (100)	30 (100)	10 (100)	75 (100)
Access to Android phone, n (%)	6 (20)	7 (23)	2 (20)	15 (20)
Able to receive and initiate phone call, n (%)	30 (100)	30 (100)	10 (100)	75 (100)
Able to read or send SMS in English or Runyankole, n (%)	25 (83)	25 (83)	7 (70)	62 (83)
Preferred medium, n (%)				
SMS text messaging	6 (20)	4 (13)	2 (20)	13 (17)
Audio messaging	5 (17)	4 (13)	3 (30)	13 (17)
Both	19 (63)	22 (73)	5 (50)	49 (65)
Preferred frequency of audio messages or SMS text messages, n (%)				
Daily	2 (7)	3 (10)	1 (10)	6 (8)
Twice weekly	15 (50)	14 (47)	4 (40)	35 (47)
Weekly	13 (43)	13 (43)	5 (50)	34 (45)
Preferred timing of the messages, n (%)				
Before 8 AM	20 (67)	17 (57)	6 (60)	48 (64)
Between 8 AM and noon	0 (0)	2 (7)	1 (10)	3 (4)
Between noon and 8 PM	5 (17)	4 (13)	1 (10)	10 (13)
Between 8 PM and midnight	5 (17)	7 (23)	2 (20)	14 (19)
Preferred length of audio call: 1-2 minutes, n (%)	30 (100)	30 (100)	10 (100)	75 (100)
Parity, median (IQR)	3 (2-4)	3 (2-5)	3 (2-4)	3 (2-4)
Household income ≥UGX 100,000 (US \$ 27.78)/month, n (%)	16 (53)	14 (47)	4 (40)	36 (48)
Antenatal care visits (≥4), n (%)	17 (57)	18 (60)	6 (60)	48 (64)
Number of people providing support, median (IQR)	10 (5-16)	12 (6-18)	10 (4-12)	10 (6-12)
Choices for skilled delivery, n (%)	18 (60)	21 (70)	7 (70)	51 (68)

^aIncludes 5 cognitive interviews.

In addition, 5 key HCPs (2 senior obstetricians—a man aged 44 years and a woman aged 39 years—and 3 experienced female midwives aged 27, 38, and 55 years) were also purposively identified and interviewed during the formative phase. Detailed preliminary findings of these interviews have been documented elsewhere [26,27].

Content Development

During data analysis, 34 topics were initially identified from the formative qualitative interviews conducted with the women and the HCPs. These topics included sexual health, mental health, family planning, comorbidities, nutrition, exercising, use of herbal medicine, ANC, pregnancy disorders, and normal labor. In all, 20 topics were considered, having been suggested by at least 50% (15/30) of the women interviewed, and 6 topics

were specifically identified by the interviewed HCPs as a critical part of the health education framework during ANC visits as per Uganda's national guidelines, making a total of 26 topics (Multimedia Appendix 2). These included birth preparedness, facility-based delivery, male partner's involvement, getting to know your health worker or HCP interaction, immunization, and caring for the baby. Appropriate content for SMS text messaging and audio messaging was developed for these key health education topics and embodied the 12 identified BCTTv1 components (Multimedia Appendix 2). All messages included both prevention and promotional information. Additional data from the formative interviews on barriers to, and facilitators of, ANC attendance and skilled delivery had been already documented [26,27].

App Prototype Design and Development

The developed SMS text messages and audio messages were uploaded onto the *eBirth* app platform in both English and Runyankole. These initial messages created the first messaging prototype. The message information could be dispensed in either English or Runyankole, depending on the recipient's preference. A total of 26 audio messages and 26 SMS text messages were developed. Each SMS text message was restricted to between 300 and 450 characters (sent as 2-3 separate but consecutive messages, each consisting of 150 characters to avoid the splitting of messages by the various small-screen phones possessed by most of the women). Each audio message was no longer than 1 minute 30 seconds.

Prototype Pilot Testing

Overview

In the qualitative interviews, 5 themes emerged related to the perceived intentions or goal of the app: attitude regarding content, wording, format, and delivery; language; frequency; length; and timing. The women identified 5 overarching app features that would maximize its impact and intended benefits. The app needed to provide (1) education and preparation, (2) motivation, encouragement, and support, (3) customization and connection, (4) easy-to-use interface, and (5) engagement and empowerment.

Education and Preparation

All participants interviewed during the pilot and the FGD indicated that the SMS text messages and audio messages received directly on their phones were beneficial, especially for first-time mothers-to-be or other inexperienced women who needed to understand and learn important information about pregnancy and childbirth. The women indicated that these messages, which had been developed with the help of a trusted and informed source, could reinforce the truth amid the limited and divergent information obtained from peers or relatives in their communities. According to these women, this timely information could support them to confidently debunk misinformation about pregnancy and childbirth over time. A pregnant participant aged 21 years stated as follows:

We don't get such accurate information from anywhere other than what we usually hear from people around us as we grow up...like every pregnancy is the same, or like it's normal or like you don't need to go to hospital which can be a big problem...with these messages coming from a real midwife, you get to learn a lot about pregnancy and childbirth and add on the little information you know and it's very good.

Most women stated that twice-weekly or weekly SMS text messages or audio messages were sufficient to communicate, educate, and prepare them through their ANC journey. Some preferred to receive messages at approximately 7 AM so as to be able to read them at the start of the day and prepare mentally during the day. Others preferred SMS text messages or audio messages to be delivered to them at approximately 8 PM when they were done with the day's busy schedules that could have distracted them and made them forget to read, internalize, share,

or discuss the messages with their significant others in real time. The women thought that this timing could offer them ample time to learn and discuss health matters with their significant others and plan the next course of action appropriately. An addition of a single midweek SMS text message reminder to the app was preferred to support their near-monthly scheduled ANC visits. The women suggested that additional SMS text message reminders sent at least once a week to their significant others (social networks) could help them to initiate or ease into a discussion on the challenges they faced and the need to go to hospital. It was indicated that stimulation of these discussions could help the women to involve their significant others in preparation for, and mobilization of, the help and support they needed for timely access to professional care during the antenatal period. A pregnant participant aged 31 years stated as follows:

I get so busy with work during the day and it's in the evening when I catch a break and read the messages very well...my husband is even around by that time so I can walk up to him and we talk about my needs together on how to like go see a doctor together just in case, so it's good that way.

A postpartum participant aged 28 years added the following:

May be once weekly early in the morning. The messages are great at the start or at the end of each week...I would love it if my mother-in-law also gets the message early like during the day so that in case I get a problem, she already knows my situation and can help me to go to hospital quickly since I do not stay with my husband all the time.

Motivation, Encouragement, and Support

The iterative interviews indicated that the women preferred messages that were both cautionary and encouraging. For example, some women suggested delivery of messages pertaining to the danger signs of a complicated pregnancy, consequences of not presenting for ANC, or incompatible sociocultural beliefs as some of the cautionary information that motivated them to actively examine themselves after reading each message. Other women indicated that cautionary or warning-based messages *sounded surreal*, had a lingering effect, and often *forced* them to quickly seek professional care as planned or as soon as these danger signs were noticed. A postpartum participant aged 34 years stated as follows:

These messages sound so surreal but, at the same time, force you to watch out in case you are not feeling well. They can wake you up and force you to go to hospital...I tried by all means to check myself all the time I received them to make sure my pain during this pregnancy was not in vain.

All the women preferred positive messages whose tone was not authoritative and that offered encouragement. Negative, rebuking messages were said to be discouraging, redundant, annoying, and emotionally draining, especially if the recipients found themselves in a situation where they had no physical or financial support to help them take action. All the women also preferred messages that were expressed in a friendly tone. A pregnant participant aged 25 years stated as follows:

That call leaves a lingering memory and a lasting impression on me from the lady who talks to me on phone, it's as if listening to someone I'm familiar with already...I mean, no one hates a good message that motivates them to do good. It's your life alright, but you need someone to encourage you to keep going like doing the right thing, you know, so you look forward to getting another one and learn from it. It's exciting.

The iterative interviews indicated that the women felt that the regular, continual information was a source of comfort and emotional support, which was provided by people they perceived as someone who cared about their well-being. Some women referred to the messaging app as a reliable “pregnancy companion” that helped them “navigate through their pregnancy.” Other women reported that the SMS text messages and audio messages gave them a sense of excitement and confidence knowing that “someone is watching over” them as they prepared to receive a new baby. An expectant participant aged 18 years stated as follows:

I get to hear encouraging messages and compare with how I feel there and then from someone who clearly understands these things and is also willing to navigate through this pregnancy journey with me...with these messages coming through every week, we know someone is watching over us and we always look forward to the next message.

A postpartum participant aged 30 years added the following:

These messages made me feel as if someone out there cared about me and my baby. Anyone would really love that because it's like you have a pregnancy companion who knows you so well and moves with you along this difficult journey. Deep down you know you will make it so you are encouraged.

Customization and Connection

Customized reminders or SMS text messages that delivered a caring message were preferred to plain default messages, reflecting users' need to be *included*, connected, and related to the program as desired. Participants reported that the addition of SMS text message reminders improved their attitude toward formal care by providing a responsive and caring *connection* to the health system. A postpartum participant aged 28 years stated as follows:

A message like, hi XX [name], your life is very important to us, we are reminding you to go for your scheduled ANC visit in time to avoid problems during your pregnancy and childbirth...It shows someone cares. It's short but such a message makes you feel good and connected to your midwife.

Most women revealed a preference for audio messages whenever possible because they provided a clearer flow of information or instructions at one go. However, many women indicated that they missed some audio messages and were unable to retrieve them, unless they called back, which would cost them. However, the system also provided them an interactive feedback mechanism that participants could use to request the caller to

repeat the current or previous messages and listen in as long as they desired using a designated number and a numeric key. The interviews revealed that missed calls tended to occur when women encountered unforeseeable events such as parties or burials where they could not answer their phones for fear of embarrassment or when their phones were in silent mode. Other occurrences included dead batteries, lost phones, and poor or no network, which affected the delivery of messages. However, unlike audio calls, SMS text messages could be delayed but eventually came through once the participants charged their phones, revived an active SIM card, or entered an area with network coverage. To resolve this, the women suggested both audio- and SMS text message-based delivery media, a call-in number to report such challenges, an option to provide an alternative SIM card number with better network, and an option to provide contacts of people within their social network who could agree to receive messages on their behalf or call them to the phone whenever they are nearby. The component of social networks was further expressed as important for improving the continuity of such messaging interventions in this setting. A postpartum participant aged 21 years observed as follows:

One can get both calls and SMS, or get a number to call you in case their phone is stolen, has problems or something else...we can even give you other numbers of the people we live with at home or nearby who can reach us and deliver the messages in case anything happens. That way, we can stay connected.

An expectant participant aged 35 years added the following:

It's good because I can call back in case I missed the call or make it repeat as many times if I did not understand anything...although I am charged some money to call in, I am sure I have got the right information and support whenever I need it.

A strong interest was expressed by the women, especially those who owned Android phones, to be able to connect and interact with other pregnant women and with their HCPs. According to these women, such a platform could help them share their ongoing challenges and lean on, and get support from, their providers or colleagues who have had the same experiences in real time. A postpartum participant aged 26 years stated as follows:

For example, I could see blood for example and be lazy to go to hospital for checkup or something, and women out there on the platform could share their lived experience and push you to go quickly...or like send you alternative contact of a midwife near you, such kind of thing. I think it's very good to get people in your condition to connect and interact with all the time.

Easy-to-Use Interface

All the women reported that they were able to receive SMS text messages or audio calls easily. When the app was modified to offer different choices of audio messages or SMS text messages to fit individual needs, SMS text messaging was mostly preferred by women who could read and write. However, some participants thought that SMS text messages became redundant

over time and instead preferred phone calls. The women observed that SMS text messaging was ideal for delivering short, concise, actionable information; therefore, a good, relevant and comprehensive message required to be split into 2-3 SMS text messages, each restricted to 150 characters, to maintain the richness of the intended information. However, the women reported that these messages sometimes came through in the wrong sequence because of network issues. Women who were unable to read preferred audio calls as an easier option because they cover the same information clearly in less than 1 minute. However, the women were also able to share SMS text messages with other people in their social networks who could read and interpret messages for them, a possible indicator that both options could be considered to reinforce information transfer and use among women in similar settings. An expectant participant aged 19 years stated as follows:

I love calls a lot because you get to listen from a real person clearly and it feels real...A phone call is also easy to receive and just listen quickly in your local language...I always take the [SMS] message to my husband or friends to read for me and so it's also good because they explain to me everything...but I get lazy and they get boring like after some time and confusing like if they are so long and one part [of the SMS] doesn't come properly or is missing...but a phone call is always clear, is brief and is hard to ignore since someone takes their time to think about you and call to encourage you.

Engagement and Empowerment

The way the message content was delivered was very important to the end users and, as such, enabled the women to continuously engage with the program. They described a need for message content that emotionally connected them to the messenger. The message content delivered by the *great messenger* was viewed as liberating and empowering because of its ability to fulfil an existing need and desire for alternative and accurate information received directly on their personal or accessible phones. As such, SMS text messages were often stored and reviewed not only as future reference material for themselves, but also for peers who needed support and with whom they shared this material. Some women were enthusiastic about audio messages, which were thought to provide more clarity, enabling them to understand and internalize the messages faster and better. Other women preferred both. The women indicated that the delivery of clear messages was important in improving understanding, engagement, and encouraging prolonged use. For example, a postpartum participant aged 31 years stated as follows:

The caller is very good. She's very clear in explaining things, she makes you understand everything, I mean such words like aka-TV [ultrasound scan]...she explains many issues in such a short time so well, and you can press and listen again or call for more clarification if you want...you get to learn a lot and teach others...It's encouraging to know that she understands us very well and I like her a lot.

A postpartum participant aged 24 years added the following:

It is very liberating to know you have this information from your musawo [health worker] sent directly on your phone at all time. I can keep the messages, read them again and can share them with my friends as much as they need it. I like it a lot.

The arrival of messages was seen as an ongoing reminder to stay alert and keep examining themselves, and this had the potential to help women to respond and make better choices as and when necessary. An expectant participant aged 27 years stated as follows:

These messages help a lot and keep you alert and informed. Every message is like a reminder for me to immediately check on a few things, like how I feel and see how I am doing right then and like quickly decide on what to do in case I am in trouble.

This information and feedback was used to modify the audio messaging and SMS text messaging app. The final messages made up a message bank for future evaluation to assess feasibility, acceptability, and preliminary efficacy to influence the initiation and use of maternity care services.

Discussion

Principal Findings

We developed a novel patient-centered SMS text messaging and audio messaging app to support women to use maternity care services in rural southwest Uganda. We found that involving end users, including providers and health care users, in developing and formulating a messaging intervention gave the women a sense of *ownership* and inclusiveness in the app development process. The women identified 5 overarching app features. The app needed to provide education and preparation; motivation, encouragement, and support; customization and connection; easy-to-use interface; and engagement and empowerment. We therefore designed and developed a novel patient-centered and customized SMS text messaging and audio messaging app to engage rural women in southwestern Uganda and communicate important targeted health-related information to support them during pregnancy. Our pilot data support a potential role for messaging apps as a complementary approach to face-to-face health education encounters, especially when the apps are well packaged and tailored to suit the end users' needs and preferences. We now plan to evaluate the intervention formally in a randomized clinical trial in the next phase of the app development process.

Prior studies have reported improved acceptability and sustainability of mHealth interventions when the apps involved individual participation and also involved end users in the initial stages of the design process [22,24,37-39]. Other scholars have documented the beneficial effect of customizing the app to improve engagement and education in the general population [33]. In these studies, individual needs, goals, expectations, intentions, and preferences needed to be addressed to enable active engagement with the app. Our app was developed with the involvement of multiple stakeholders, including end users, HCPs, and health and technology experts. Our analysis reinforces these findings, particularly among women who

expressed a great desire for relevant and customized mHealth apps and content that educates, prepares, motivates, encourages, and supports them during pregnancy. We developed customized messages tailored to these end-user preferences for this intervention and opted for a standardized automated messaging program to ensure delivery of regular health information directly to users' phones. This approach helped women to stay connected and motivated, and they would refer to the intervention as a "pregnancy companion."

We used an iterative approach to develop, test, and deliver a user-driven app that was considered appropriate for a largely non-Android user community to improve utility, usability, desirability, and uptake. This was in line with previous studies that found SMS text messaging language, medium of message delivery, experience with similar technology, phone type, and characteristics to be critical in designing and delivering a culturally appropriate mHealth program that is consistent with the way people are already using the promoted technology [37,39]. The creation of an intervention that was compatible with local mobile phone types and the provision of varying delivery media for both literate and illiterate individuals to receive or access credible information support from the system, keep it for future reference, or share it with family and significant others directly on their regular phones seemed to shape the users' perception of the intervention's usefulness. We will next assess the impact of this iterative design approach through a planned randomized clinical trial and, if successful, through programmatic implementation to evaluate its influence on health outcomes and adoption in practice.

The end users identified key technical and design preferences and challenges of the mHealth app related to the timing and frequency of messaging, unreliable phone batteries, network issues, and lost phones. The women often chose weekly SMS text message reminders, possibly because scheduled ANC visits are near-monthly, and 2 promotional or cautionary messages at the beginning and end of each week as sufficient without causing unnecessary burden. According to previous scholars, this establishment of different preferences regarding timing and frequency helps to avoid unnecessary repetition, technology fatigue, and boredom, making the messaging intervention an acceptable tool to deliver health promotion content [40]. Shaw et al [40], however, observed that sending messages at the same time of the day could reduce the value that participants accord the messages and reduce the frequency of responses. The timing of messages at the start and end of the day were therefore varied over the course of the program but fixed at the beginning and end of the week as preferred to mitigate network and battery issues. The women in our pilot interviews additionally suggested providing an option to register both numbers of a dual SIM card mobile phone or to provide other contact numbers in their social networks, with consent, to receive reminders and messages as an alternative means to improve network connectivity, continuity, and reliability, as well as minimize the issue of missed messages because of network, phone, and battery issues apparent in this setting. In line with previous studies [23], the involvement of significant others in the pilot seemed to facilitate messaging continuity and encourage women to initiate important

health discussions, enabling mobilization of needed resources and support for timely access to maternity care.

Our study included a number of strengths. We involved both providers and health care users in developing a user-friendly, culture-consistent, patient-centered automated SMS text messaging and audio messaging app to stimulate, encourage, and support rural women in southwestern Uganda to use maternity care services. This study also documents women's expectations, experiences, perceptions, and choices of an mHealth-based technology that would benefit and support them in their local communities, subject to the standard limitations of network challenges as well as mobile phone ownership and type in the region. We developed a total of 26 audio messages (based on the 26 identified ANC topics), delivered in messages restricted to 150 characters per message to suit the different phone types accessed by women in this community. This stepped and multidisciplinary approach can inform the design and implementation of other novel patient-centered interventions that aim to reinforce *ownership*, engagement, inclusiveness, and uptake in a program operating in local communities. Our study also demonstrates a potential approach that can be used to complement face-to-face education encounters by communicating important targeted health-related information that is beneficial and offers informational and emotional social support to rural women through a novel automated audio messaging and SMS text messaging app. We used well-established conceptual frameworks to characterize and develop message content for this app, helping to make the findings more grounded, acceptable, meaningful, and generalizable.

This study also included some limitations. Our approach of using conceptual frameworks may have limited considerations of key variables, which may have influenced our direction and content of interest. There is therefore a need to assess how these specific variables differ under different settings or circumstances. Most of the people in our study setting are from a less affluent or less educated background, and they are not smartphone users, which imposes limits on internet access despite improved internet penetration through local mobile phone companies. As such, the messaging content and delivery medium were developed to suit current phone access and characteristics in similar settings and thus might not be generalizable to other settings with higher literacy or smartphone use. We were not able to develop a platform to enable direct and automated feedback and affirmation on self-goal attainment targets, an approach likely to motivate long-term use [33]. However, the ability of our app to engage with social networks and HCPs was anticipated to help women continuously share their experiences concerning their milestones, challenges, and goal attainments. Sharing such successes for others to acknowledge and learn from their own accomplishments has been documented to exceptionally motivate app users [33]. We did not assess for feasibility and acceptability of the messages in this particular work. The final version of the app is currently undergoing a pilot clinical trial to document feasibility, acceptability, and preliminary efficacy.

Conclusions

Our study describes a process for developing and testing a novel user-friendly, patient-centered mHealth-based messaging app suitable for rural women in southwestern Uganda. We have demonstrated an iterative approach with diffuse stakeholders to develop a customized, automated, patient-centered audio messaging and SMS text messaging app designed to communicate important targeted health-related information and support rural pregnant women in southwestern Uganda.

Involving both HCPs and end users in developing and formulating the mHealth intervention allowed us to tailor the intervention characteristics to the women's preferences, thus giving them a sense of ownership and inclusiveness in the program. This approach supports a potential role for messaging apps as a complementary approach to face-to-face health education encounters. Our next step is to evaluate the intervention in a pilot clinical trial to assess its larger-scale feasibility, acceptability, and ability to affect health outcomes.

Acknowledgments

The authors would like to acknowledge the staff of Mbarara Regional Referral Hospital, the women who participated in the interviews, and *iStreams* for their contribution. Many thanks to research assistants Edgar Agaba and Eunice Atuheirwe. The research reported in this study was supported by the Fogarty International Center of the National Institutes of Health under award number K43TW011004.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guides.

[[PDF File \(Adobe PDF File\), 550 KB - formative_v5i11e29214_app1.pdf](#)]

Multimedia Appendix 2

Message content.

[[PDF File \(Adobe PDF File\), 658 KB - formative_v5i11e29214_app2.pdf](#)]

References

1. Campbell OM, Graham WJ. Strategies for reducing maternal mortality: getting on with what works. *Lancet* 2006 Oct;368(9543):1284-1299. [doi: [10.1016/s0140-6736\(06\)69381-1](#)]
2. Campbell OM, Benova L, MacLeod D, Baggaley RF, Rodrigues LC, Hanson K, et al. Family planning, antenatal and delivery care: cross-sectional survey evidence on levels of coverage and inequalities by public and private sector in 57 low- and middle-income countries. *Trop Med Int Health* 2016 Apr 07;21(4):486-503 [FREE Full text] [doi: [10.1111/tmi.12681](#)] [Medline: [26892335](#)]
3. World Health Organization. WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience. Geneva: World Health Organization; 2016:1-196.
4. Carroli G, Rooney C, Villar J. How effective is antenatal care in preventing maternal mortality and serious morbidity? An overview of the evidence. *Paediatr Perinat Epidemiol* 2001 Jan;15 Suppl 1(s1):1-42. [doi: [10.1046/j.1365-3016.2001.0150s1001.x](#)] [Medline: [11243499](#)]
5. Sondaal SF, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mhealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. *PLoS One* 2016 May 4;11(5):e0154664 [FREE Full text] [doi: [10.1371/journal.pone.0154664](#)] [Medline: [27144393](#)]
6. Uganda Demographic and Health Survey 2016. Uganda Bureau of Statistics - UBOS and ICF. 2018. URL: <https://dhsprogram.com/publications/publication-fr333-dhs-final-reports.cfm> [accessed 2021-11-08]
7. Thomson G, Balaam M, Hymers K. Building social capital through breastfeeding peer support: insights from an evaluation of a voluntary breastfeeding peer support service in North-West England. *Int Breastfeed J* 2015 Apr 2;10(1):15 [FREE Full text] [doi: [10.1186/s13006-015-0039-4](#)] [Medline: [25897318](#)]
8. Bootsri W, Taneepanichskul S. Effectiveness of experiential learning with empowerment strategies and social support from grandmothers on breastfeeding among Thai adolescent mothers. *Int Breastfeed J* 2017 Aug 22;12(1):37 [FREE Full text] [doi: [10.1186/s13006-017-0128-7](#)] [Medline: [28852416](#)]
9. Gonzalez JS, Penedo FJ, Antoni MH, Durán RE, McPherson-Baker S, Ironson G, et al. Social support, positive states of mind, and HIV treatment adherence in men and women living with HIV/AIDS. *Health Psychol* 2004 Jul;23(4):413-418. [doi: [10.1037/0278-6133.23.4.413](#)] [Medline: [15264978](#)]
10. Edwards LV. Perceived social support and HIV/AIDS medication adherence among African American women. *Qual Health Res* 2006 May 01;16(5):679-691. [doi: [10.1177/1049732305281597](#)] [Medline: [16611972](#)]

11. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. *J Med Internet Res* 2007 Apr 27;9(2):e7 [FREE Full text] [doi: [10.2196/jmir.9.2.e7](https://doi.org/10.2196/jmir.9.2.e7)] [Medline: [17478409](https://pubmed.ncbi.nlm.nih.gov/17478409/)]
12. Moore MJ, Soderquist J, Werch C. Feasibility and efficacy of a binge drinking prevention intervention for college students delivered via the internet versus postal mail. *J Am Coll Health* 2005 Jul;54(1):38-44. [doi: [10.3200/jach.54.1.38-44](https://doi.org/10.3200/jach.54.1.38-44)]
13. Hurling R, Fairley BW, Dias MB. Internet-based exercise intervention systems: are more interactive designs better? *Psychol Health* 2006 Dec;21(6):757-772. [doi: [10.1080/14768320600603257](https://doi.org/10.1080/14768320600603257)]
14. Mobile cellular subscriptions (per 100 people) - Uganda. World Bank. 2016. URL: https://data.worldbank.org/indicator/IT.CEL.SETS.P2?locations=UG&most_recent_year_desc=false [accessed 2021-11-09]
15. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? *PLoS Med* 2013 Feb 12;10(2):e1001382 [FREE Full text] [doi: [10.1371/journal.pmed.1001382](https://doi.org/10.1371/journal.pmed.1001382)] [Medline: [23424286](https://pubmed.ncbi.nlm.nih.gov/23424286/)]
16. Nurmatov UB, Lee SH, Nwaru BI, Mukherjee M, Grant L, Pagliari C. The effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: protocol for a systematic review and meta-analysis. *J Glob Health* 2014 Jun;4(1):010407 [FREE Full text] [doi: [10.7189/jogh.04.010407](https://doi.org/10.7189/jogh.04.010407)] [Medline: [24976966](https://pubmed.ncbi.nlm.nih.gov/24976966/)]
17. Long H, Bartlett YK, Farmer AJ, French DP. Identifying brief message content for interventions delivered via mobile devices to improve medication adherence in people with type 2 diabetes mellitus: a rapid systematic review. *J Med Internet Res* 2019 Jan 09;21(1):e10421 [FREE Full text] [doi: [10.2196/10421](https://doi.org/10.2196/10421)] [Medline: [30626562](https://pubmed.ncbi.nlm.nih.gov/30626562/)]
18. Higgs ES, Goldberg AB, Labrique AB, Cook SH, Schmid C, Cole CF, et al. Understanding the role of mHealth and other media interventions for behavior change to enhance child survival and development in low- and middle-income countries: an evidence review. *J Health Commun* 2014 Sep 10;19 Suppl 1(sup1):164-189 [FREE Full text] [doi: [10.1080/10810730.2014.929763](https://doi.org/10.1080/10810730.2014.929763)] [Medline: [25207452](https://pubmed.ncbi.nlm.nih.gov/25207452/)]
19. Fedha T. Impact of mobile telephone on maternal health service care: a case of Njoro division. *Open J Prevent Med* 2014;04(05):365-376. [doi: [10.4236/ojpm.2014.45044](https://doi.org/10.4236/ojpm.2014.45044)]
20. Karen G, Rimer BK, Viswanath K. Health Behavior and Health Education: Theory, Research, and Practice 4th Edition. Hoboken, New Jersey, United States: John Wiley & Sons; 2008:1-592.
21. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. *Glob Health Sci Pract* 2013 Aug 06;1(2):160-171. [doi: [10.9745/ghsp-d-13-00031](https://doi.org/10.9745/ghsp-d-13-00031)]
22. Lotto M, Strieder AP, Aguirre PE, Oliveira TM, Machado MA, Rios D, et al. Parental-oriented educational mobile messages to aid in the control of early childhood caries in low socioeconomic children: a randomized controlled trial. *J Dent* 2020 Oct;101:103456 [FREE Full text] [doi: [10.1016/j.jdent.2020.103456](https://doi.org/10.1016/j.jdent.2020.103456)] [Medline: [32827598](https://pubmed.ncbi.nlm.nih.gov/32827598/)]
23. Atukunda EC, Musiimenta A, Musinguzi N, Wyatt MA, Ashaba J, Ware NC, et al. Understanding patterns of social support and their relationship to an art adherence intervention among adults in rural Southwestern Uganda. *AIDS Behav* 2017 Feb 26;21(2):428-440 [FREE Full text] [doi: [10.1007/s10461-016-1559-7](https://doi.org/10.1007/s10461-016-1559-7)] [Medline: [27671479](https://pubmed.ncbi.nlm.nih.gov/27671479/)]
24. 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/)]
25. Ministry of Health. Community Health Extension Workers Strategy in Uganda (2015/16- 2019/20). Kampala, Uganda: Ministry of Health; 2019.
26. Atukunda EC, Mugenyi GR, Obua C, Musiimenta A, Agaba E, Najjuma JN, et al. Women's choice to deliver at home: understanding the psychosocial and cultural factors influencing birthing choices for unskilled home delivery among women in Southwestern Uganda. *J Pregnancy* 2020 Jun 03;2020:6596394 [FREE Full text] [doi: [10.1155/2020/6596394](https://doi.org/10.1155/2020/6596394)] [Medline: [32566299](https://pubmed.ncbi.nlm.nih.gov/32566299/)]
27. Atukunda EC, Mugenyi GR, Obua C, Musiimenta A, Najjuma JN, Agaba E, et al. When women deliver at home without a skilled birth attendant: a qualitative study on the role of health care systems in the increasing home births among rural women in Southwestern Uganda. *Int J Womens Health* 2020 May; Volume 12:423-434. [doi: [10.2147/ijwh.s248240](https://doi.org/10.2147/ijwh.s248240)]
28. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010 Feb;43(1):159-172 [FREE Full text] [doi: [10.1016/j.jbi.2009.07.002](https://doi.org/10.1016/j.jbi.2009.07.002)] [Medline: [19615467](https://pubmed.ncbi.nlm.nih.gov/19615467/)]
29. van Eijk D, van Kuijk J, Hoolhorst F, Kim C, Harkema C, Dorrestijn S. Design for usability; practice-oriented research for user-centered product design. *Work* 2012;41 Suppl 1:1008-1015. [doi: [10.3233/WOR-2012-1010-1008](https://doi.org/10.3233/WOR-2012-1010-1008)] [Medline: [22316853](https://pubmed.ncbi.nlm.nih.gov/22316853/)]
30. Uganda Clinical Guidelines: Antenatal Care (ANC). Ministry of Health: Kampala, Uganda. 2016. URL: https://health.go.ug/sites/default/files/Uganda%20Clinical%20Guidelines%202016_FINAL.pdf [accessed 2021-11-09]
31. 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. [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/)]
32. Liang L, Wiens MO, Lubega P, Spillman I, Mugisha S. A Locally Developed Electronic Health Platform in Uganda: Development and Implementation of Stre@mline. *JMIR Form Res* 2018 Sep 24;2(2):e20 [FREE Full text] [doi: [10.2196/formative.9658](https://doi.org/10.2196/formative.9658)] [Medline: [30684419](https://pubmed.ncbi.nlm.nih.gov/30684419/)]

33. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. *JMIR Mhealth Uhealth* 2017 Sep 26;5(9):e141 [[FREE Full text](#)] [doi: [10.2196/mhealth.7136](https://doi.org/10.2196/mhealth.7136)] [Medline: [28951378](#)]
34. 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](#)]
35. Krueger RA, Casey MA. *Focus Groups: A Practical Guide for Applied Research*. Thousand Oaks, CA: SAGE Publications, Inc; 2008:1-240.
36. Bronfenbrenner U. Toward an experimental ecology of human development. *Am Psychol* 1977;32(7):513-531. [doi: [10.1037/0003-066x.32.7.513](https://doi.org/10.1037/0003-066x.32.7.513)]
37. Campbell JI, Aturinda I, Mwesigwa E, Burns B, Santorino D, Haberer JE, et al. The Technology Acceptance Model for Resource-Limited Settings (TAM-RLS): a novel framework for mobile health interventions targeted to low-literacy end-users in resource-limited settings. *AIDS Behav* 2017 Nov 18;21(11):3129-3140 [[FREE Full text](#)] [doi: [10.1007/s10461-017-1765-y](https://doi.org/10.1007/s10461-017-1765-y)] [Medline: [28421356](#)]
38. Ybarra ML, Holtrop JS, Bağcı Bosi AT, Emri S. Design considerations in developing a text messaging program aimed at smoking cessation. *J Med Internet Res* 2012 Jul 24;14(4):e103 [[FREE Full text](#)] [doi: [10.2196/jmir.2061](https://doi.org/10.2196/jmir.2061)] [Medline: [22832182](#)]
39. Dick JJ, Nundy S, Solomon MC, Bishop KN, Chin MH, Peek ME. Feasibility and usability of a text message-based program for diabetes self-management in an urban African-American population. *J Diabetes Sci Technol* 2011 Sep 01;5(5):1246-1254 [[FREE Full text](#)] [doi: [10.1177/193229681100500534](https://doi.org/10.1177/193229681100500534)] [Medline: [22027326](#)]
40. Shaw RJ, Bosworth HB, Hess JC, Silva SG, Lipkus IM, Davis LL, et al. Development of a theoretically driven mHealth text messaging application for sustaining recent weight loss. *JMIR Mhealth Uhealth* 2013 May 07;1(1):e5 [[FREE Full text](#)] [doi: [10.2196/mhealth.2343](https://doi.org/10.2196/mhealth.2343)] [Medline: [25100678](#)]

Abbreviations

- ANC:** antenatal care
BCTTv1: Behavior Change Technique Taxonomy version 1
FGD: focus group discussion
HCP: health care provider
mHealth: mobile health
MRRH: Mbarara Regional Referral Hospital
VHT: village health team

Edited by G Eysenbach; submitted 30.03.21; peer-reviewed by M Lotto, J Zlamal, E Nabovati, C Reis, M Ofori; comments to author 08.09.21; revised version received 05.10.21; accepted 27.10.21; published 25.11.21.

Please cite as:

Atukunda EC, Matthews LT, Musiimenta A, Mugenyi GR, Mugisha S, Ware NC, Obua C, Siedner MJ
mHealth-Based Health Promotion Intervention to Improve Use of Maternity Care Services Among Women in Rural Southwestern Uganda: Iterative Development Study
JMIR Form Res 2021;5(11):e29214
URL: <https://formative.jmir.org/2021/11/e29214>
doi: [10.2196/29214](https://doi.org/10.2196/29214)
PMID: [34842541](https://pubmed.ncbi.nlm.nih.gov/34842541/)

©Esther Cathyln Atukunda, Lynn T Matthews, Angella Musiimenta, Godfrey Rwambuka Mugenyi, Samuel Mugisha, Norma C Ware, Celestino Obua, Mark J Siedner. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 25.11.2021. 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

Patients' Use of Mobile Health for Self-management of Knee Osteoarthritis: Results of a 6-Week Pilot Study

Brittany Shewchuk¹, BSc, MSc; Lee A Green^{2,3}, MPH, MD; Tanya Barber³, MA; Jean Miller⁴, PhD; Sylvia Teare⁴, MED; Denise Campbell-Scherer^{5,6}, MD, PhD; Kelly J Mrklas^{1,7}, BSc, MSc; Linda C Li⁸, BSc, PT, MSc, PhD; Nancy Marlett^{9,10}, PhD; Tracy Wasylak^{11,12}, MSc, CHE; Elena Lopatina¹³, MD, PhD; Deirdre McCaughey¹, MBA, PhD; Deborah A Marshall¹, PhD

¹Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

²Department of Family Medicine, University of Alberta, Edmonton, AB, Canada

³Enhancing Alberta Primary Care Research Networks, Department of Family Medicine, University of Alberta, Edmonton, AB, Canada

⁴O'Brien Institute for Public Health, University of Calgary, Calgary, AB, Canada

⁵Department of Family Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

⁶Office of Lifelong Learning and Physician Learning Program, Faculty of Medicine and Dentistry, Edmonton, AB, Canada

⁷Strategic Clinical Networks, System Innovation and Programs, Alberta Health Services, Calgary, AB, Canada

⁸Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada

⁹Patient and Community Engagement Research Unit, O'Brien Institute for Public Health, Calgary, AB, Canada

¹⁰Community Rehabilitation and Disability Studies, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

¹¹Strategic Clinical Networks, Alberta Health Services, Calgary, AB, Canada

¹²Faculty of Nursing, University of Calgary, Calgary, AB, Canada

¹³Department of Community Health Sciences, Faculty of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:

Deborah A Marshall, PhD

Department of Community Health Sciences

Cumming School of Medicine

University of Calgary

3C58 Health Research Innovation Centre, 3230 Hospital Drive

Calgary, AB, T2N 4Z6

Canada

Phone: 1 403 210 6377

Email: damarsha@ucalgary.ca

Abstract

Background: In a previous study, a prototype mobile health (mHealth) app was co-designed with patients, family physicians, and researchers to enhance self-management and optimize conservative management for patients with mild to moderate knee osteoarthritis (OA).

Objective: This study aims to evaluate the overall usability, quality, and effectiveness of the mHealth app prototype for aiding knee OA self-management from the perspectives of patients with OA and health care providers (HCPs).

Methods: Using methods triangulation of qualitative and quantitative data, we conducted a pilot evaluation of an mHealth app prototype that was codeveloped with patients and HCPs. We recruited adult patients aged ≥ 20 years with early knee OA ($n=18$) who experienced knee pain on most days of the month at any time in the past and HCPs ($n=7$) to participate. In the qualitative assessment, patient and HCP perspectives were elicited on the likeability and usefulness of app features and functionalities and the perceived impact of the app on patient-HCP communication. The quantitative assessment involved evaluating the app using usability, quality, and effectiveness metrics. Patient baseline assessments included a semistructured interview and survey to gather demographics and assess the quality of life (European Quality-of-Life 5-Dimension 5-Level Questionnaire [EQ-5D-5L]) and patient activation (patient activation measure [PAM]). Following the 6-week usability trial period, a follow-up survey assessed patients' perceptions of app usability and quality and longitudinal changes in quality of life and patient activation. Semistructured interviews and surveys were also conducted with HCPs ($n=7$) at baseline to evaluate the usability and quality of the app prototype.

Results: Interviews with patients and HCPs revealed overall positive impressions of the app prototype features and functionalities related to likeability and usefulness. Between the baseline and follow-up patient assessments, the mean EQ-5D-5L scores improved from 0.77 to 0.67 ($P=.04$), and PAM scores increased from 80.4 to 87.9 ($P=.01$). Following the 6-week evaluation, patients reported a mean System Usability Scale (SUS) score of 57.8, indicating marginal acceptability according to SUS cutoffs. The mean number of goals set during the usability period was 2.47 (SD 3.08), and the mean number of activities completed for knee OA self-management during the study period was 22.2 (SD 17.8). Spearman rank correlation (r_s) calculations revealed that the follow-up PAM scores were weakly correlated ($r_s=-0.32$) with the number of goals achieved and the number ($r_s=0.19$) of activities performed during the 6-week usability period. HCPs reported a mean SUS score of 39.1, indicating unacceptable usability.

Conclusions: This evidence-based and patient-centered app prototype represents a potential use of mHealth for improving outcomes and enhancing conservative care by promoting patient activation and patient-HCP communication regarding OA management. However, future iterations of the app prototype are required to address the limitations related to usability and quality.

(*JMIR Form Res* 2021;5(11):e30495) doi:[10.2196/30495](https://doi.org/10.2196/30495)

KEYWORDS

mobile health; mHealth; app; self-management; osteoarthritis

Introduction

Background

Osteoarthritis (OA) is the most common form of arthritis and a leading cause of disability worldwide [1]. Current management strategies for knee OA largely target symptom control with the primary aim of reducing pain, improving function, and delaying the need for surgical intervention [2]. There are currently no disease-modifying therapies for OA, and existing pharmacological treatments for targeting symptoms are largely ineffective [3,4]. Nonpharmacological interventions such as exercise, weight management, and patient education have been shown to be effective in reducing OA-related pain and disability [5,6].

OA Self-management

The latest clinical guidelines for OA recommend self-management, including health education and goal setting, as the core treatment alongside nonpharmacological and pharmacological treatments [7,8] when used as a supplement to medical care [9]. In the context of chronic disease management, goal setting and action planning have been found to improve self-efficacy, encourage behavioral change, and improve health outcomes [10-12]. Educational interventions, when delivered alongside exercise and weight loss programs, have also been shown to lead to better treatment adherence, reductions in pain, better self-management, and improved quality of life [13]. OA self-management tools can have an important impact on improving the delivery of interventional programs targeting patient education [14] and behavioral modification [15] and can also influence the effectiveness of patient–health care provider (HCP) communication and shared decision-making [9,16]. As such, acceptance and adoption of mobile health (mHealth) strategies for self-management can help improve health outcomes, reduce costs to the health care system, and encourage patients to take a more active role in improving their health.

mHealth Apps

mHealth apps have shown promise in supporting patient self-management of health conditions, especially for chronic diseases that require long-term care and maintenance [9]. However, barriers to the adoption and use of mHealth technologies have been identified. First, the lack of user-centered designs integrating patients' needs and preferences intercepts the adoption of and sustained engagement with such mHealth technologies [9,17]. Second, there has been a lack of formal mHealth evaluations to date [18,19], raising concerns about the safety and effectiveness of mHealth technologies [20,21]. Our understanding of the effectiveness of mHealth apps in supporting self-management and their impact on patient-reported outcomes in adults with knee OA remains in its early stages. Thus, there is both a need to fill the gap in the availability of mHealth apps for knee OA management and provide a framework to evaluate mHealth apps targeting OA self-management driven by the priorities and feedback of target users.

Objectives

In our previous work, we explored the perspectives of end users on an mHealth app for knee OA self-management, leading to the co-design of an app prototype aimed at facilitating self-management and improving patient–physician communication [22,23]. The resulting app reflected a consensus of patient and HCP priority functional requirements, achieved through co-design. In this study, the overarching objective is to evaluate the app from both patient and HCP perspectives using qualitative and quantitative assessments. The primary aims of the 2-part patient evaluation are to assess the app by its (1) overall usability and quality, (2) ability to improve patient self-management behavior (goal setting and activity completion), and (3) effectiveness in improving quality of life, patient activation, and patient-HCP communication. The primary aim of the 1-time HCP evaluation is to assess the overall usability, quality, and perceived impact on patient-HCP communication.

Methods

Overview

We used a combination of qualitative and quantitative methods and triangulated data for increased rigor to evaluate the app features and functionalities from patient and HCP perspectives. Details on the main features and functionalities included in the app, which were determined in preceding co-design sessions, are summarized in [Textbox 1](#). The app was assessed at baseline

by patients and HCPs using semistructured interviews and surveys, followed by a 6-week usability period and a final follow-up survey evaluation for patients only. This study was reviewed and approved by the University of Calgary research ethics board (REB16-1372). Privacy and confidentiality of data were maintained throughout all phases of the study. All personal identifying information was stripped from interview documents and recruitment materials, and all the data were deidentified and stored in a password-protected electronic file on secure servers at the University of Calgary.

Textbox 1. Functionalities associated with each feature of the final mobile health app prototype.

<p>Symptom tracking (dashboard)</p> <ul style="list-style-type: none"> • Symptoms are tracked on the following dimensions: pain, stiffness, and functional limitation. • Data on symptoms gathered from the patients were evaluated according to a threshold-based approach based on the Western Ontario and McMaster Universities Osteoarthritis Index criteria. <p>Goals</p> <ul style="list-style-type: none"> • Patients prospectively identify goals from a set list of categories (exercise and activity, pain reduction, and weight loss). • Goals are further delineated into customized activities using the specific, measurable, attainable, relevant, and time-based goal-setting schema. • Goals and activities can be tracked over a limited or continuous period. <p>Activities</p> <ul style="list-style-type: none"> • Activity is tracked on the following 2 dimensions: (1) activities related to goal setting (eg, linked to favorite activity, such as golf or biking) and (2) exercise (evidence-based recommendations from a physiotherapist or other health care provider) • Activity categories provided were aerobic activity, aquatic activity, muscle strengthening, and others. <p>Red flags</p> <ul style="list-style-type: none"> • Red flags can be identified by patients in a journal-like manner, including any experienced symptom or activity-related difficulties that users may wish to discuss with their health care providers. • Red flag categories provided were infection, trauma (eg, fracture), persistent inflammation, warmth, swelling or persistent pain, low mood, activity avoidance, or others. <p>Resources</p> <ul style="list-style-type: none"> • The summary of evidence-based self-management resources for knee osteoarthritis self-management was defined by 3 main headings: (1) information for self-management, (2) exercise therapies, and (3) guidance for goal setting.
--

Patient Evaluation

Recruitment

Patients were recruited in partnership with the Arthritis Alliance of Canada and the Alberta Strategy for Patient-Oriented Research Support Unit (16/18, 89%) to capture a range of patient perspectives. We also recruited those who were included in preceding co-design sessions and expressed interest in participating (2/18, 11%). Adults aged ≥ 20 years with early knee OA and who experienced knee pain on most days of the month at any time in the past were eligible to participate. Patient participants completed a baseline (week 0) semistructured interview and survey evaluation, followed by a 6-week usability period and final (week 6) follow-up survey evaluation.

Qualitative Evaluation

Patients were provided with a weblink to the app 1 week before the launch of the baseline evaluation to allow them sufficient time to familiarize themselves with the main app features

(dashboard and symptom tracking, goals, activities, red flags, and resources). At baseline, research team members (JM and ST) from the Patient and Community Engagement Research (PaCER) facilitated semistructured *talk-out-loud* interviews with patient participants, which served as an orientation to the app features and functionalities. The interview guide for the qualitative evaluation can be found in [Multimedia Appendix 1](#). Interviews were conducted via telephone or in person depending on the patients' preferences and were audio recorded for subsequent validation. The main objective of the baseline interview was to elicit patients' initial impressions of the app related to user experience. Patient perspectives on the 5 main app features were elicited on the following parameters: (1) usefulness, (2) ease of use, (3) contribution to OA knowledge, (4) self-management potential, and (5) perceptions of the impact on patient-HCP communication. With the goal of summarizing the patient experience when using different app features [24], PaCER researchers subsequently used descriptive analysis to generate a compilation of descriptive statements that reflected

the scope of responses for each of the 5 main app features. The descriptive statements were summarized using the following headings: likeability, usefulness, areas lacking and suggestions, and usability.

Quantitative Evaluation

Patient-reported measures related to app evaluation were collected using surveys administered at baseline and follow-up. The baseline survey collected information on patient demographics, history of OA symptoms and risk factors, quality of life, self-management behavior, and patient activation.

This baseline assessment was followed by a 6-week pilot trial in which patient metadata, including number of goals set and number of activities, were captured from direct inputs into the app and frequency of daily use. The follow-up survey collected information on perceived app usability and quality and reassessed postpilot quality of life and patient activation.

HCP Evaluation

Recruitment

HCPs (n=7) were recruited through our primary care research partners, Enhancing Alberta Primary Care Research Networks and Accelerating Change Transformation Team (ACTT), using purposive sampling [25]. Enhancing Alberta Primary Care Research Networks and ACTT invited HCPs who represented the early majority, defined as the first sizable group of providers to adopt an innovation after it has been established by early adopters [26]. Participating HCPs included those who partook in the preceding co-design sessions (5/7, 71%) and those recruited via existing relationships with practices and primary care networks (2/7, 29%).

Qualitative and Quantitative Evaluations

The HCP assessment involved a 1-time evaluation comprising a semistructured interview and survey (see [Multimedia Appendix 2](#) for the interview guide). During the interview component, 1 researcher from ACTT explored the app with the HCP and, using the *think-aloud* method [27,28], elicited their views on the clinical utility of the app, with emphasis on the perceived impact on patient-HCP communication. Interviews took place over the phone at a time of the HCP's choosing, were approximately 45 minutes in length, and were audio recorded for subsequent analysis. The audio recordings were not transcribed but used later to validate the notes taken during the call. At the time of the interview, both the researcher and the HCP had access to the app to navigate its functions together.

HCPs also completed a questionnaire to measure their perceptions of the app's usability and quality.

Evaluation Measures

Surveys were administered to patients at baseline and follow-up to evaluate changes in patient-reported outcomes over the course of the 6-week pilot trial. At the baseline patient evaluation, data were captured on demographics, knee symptoms, OA risk factors, quality of life, and patient activation. Quality of life was assessed using the European Quality-of-Life 5-Dimension 5-Level Questionnaire (EQ-5D-5L), a preference-based measure for describing and evaluating health, covering 5 dimensions (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) scored on a 5-point Likert scale [29]. The 10-item Patient Activation Measure (PAM-10) was administered to patients at week 0 and week 6 to provide a measure of patient activation, defined as a person's knowledge, skills, and confidence related to managing their own health [30]. The PAM-10 was designed to minimize the response burden from its 13-item successor, the PAM-13, and has been reported to have comparable levels of consistency and reliability with the PAM-13 [31]. The PAM provides a measure of a patient's engagement in self-management of their disease and thus enables change in patient activation to be monitored between baseline and follow-up evaluations.

At the patient follow-up evaluation, the EQ-5D-5L and PAM tools were re-administered along with additional measures assessing perceived app usability and quality. App usability was measured using the System Usability Scale (SUS), a well-established psychometric tool used worldwide with high levels of reported validity and reliability [32]. App quality was assessed using the App Chronic Disease Checklist (ACDC), an expert opinion-based checklist developed to evaluate the usability of chronic disease apps for monitoring, self-management, and behavioral change [33]. Finally, 2 sections (app subjective quality and app-specific quality) from the Mobile App Rating Scale (MARS) were used to provide an objective rating of perceived app quality [34], with each question measured on a 5-point Likert scale. However, it should be noted that not all sections from the ACDC and MARS tools were included in the assessment, as they included dimensions beyond the scope of the app prototype features (eg, gamification).

During the 1-time evaluation with HCPs, the SUS and MARS measures were used to assess perceived app usability and quality, respectively, from the provider perspective. Further details on each survey instrument used in the pilot testing evaluation are included in [Tables 1](#) and [2](#).

Table 1. Description of measures used for qualitative evaluation of the app prototype throughout the 6-week pilot trial.

Measurement	Patients		HCPs ^a at baseline	Method of elicitation
	Baseline	Follow-up		
Likeability, usefulness, areas lacking or suggestions, and usability	✓		✓	Semistructured <i>talk-out-loud</i> interview; see Multimedia Appendices 1 and 2 for patient and HCP qualitative interview guides, respectively, and Multimedia Appendices 3 and 4 for qualitative reports from qualitative interviews for patients and HCPs, respectively
Enhancing patient-HCP communication	✓		✓	Semistructured <i>talk-out-loud</i> interview; see Multimedia Appendices 1 and 2 for patient and HCP qualitative interview guides, respectively, and Multimedia Appendices 3 and 4 for qualitative reports from qualitative interviews for patients and HCPs, respectively
Increasing OA ^b knowledge	✓			Semistructured <i>talk-out-loud</i> interview; see Multimedia Appendices 1 and 2 for patient and HCP qualitative interview guides, respectively, and Multimedia Appendices 3 and 4 for qualitative reports from qualitative interviews for patients and HCPs, respectively
Improving OA self-management	✓			Semistructured <i>talk-out-loud</i> interview; see Multimedia Appendices 1 and 2 for patient and HCP qualitative interview guides, respectively, and Multimedia Appendices 3 and 4 for qualitative reports from qualitative interviews for patients and HCPs, respectively

^aHCP: health care provider.

^bOA: osteoarthritis.

Table 2. Description of measures used for quantitative evaluation of the app prototype throughout the 6-week pilot trial.

Measurement and instrument	Patients		HCPs ^a at baseline	Validated tool	Scoring methods	
	Baseline	Follow-up			Methods	Interpretation
Quality of life						
EQ-5D-5L ^b	✓ ^c	✓		✓	The EQ-5D-5L measures HRQoL ^d on 5 dimensions (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression). The health state of each participant is determined (series of 5 numbers corresponding to individual selections for each dimension) and then converted to an index score using a value set generated by and validated for the Canadian population. ^e	Index scores range from -0.15 to 0.95 using the Canadian value set, where low scores correspond to higher HRQoL, and high scores correspond to lower HRQoL; MID ^f , which is the minimum important change in EQ-5D-5L, scores are determined for specific patient populations and used to interpret EQ-5D-5L scores (MID for degenerative knee population=0.20).
Patient activation						
PAM-10 ^g	✓	✓		✓	A raw PAM score is calculated by summing responses for all 10 PAM questions for each respondent (scored on a 4-point Likert scale, where 1=nonactivated and 4=highly activated) and dividing the sum by the number of questions completed; mean PAM scores are converted to activation scores (scale from 0 to 100) using an empirically derived calibration table.	<ul style="list-style-type: none"> • PAM score ≤47.0: people tend to be overwhelmed and unprepared to play an active role; they are predisposed to be passive recipients of care. • PAM score 47.1-55.1: individuals lack knowledge and confidence for self-management. • PAM score 55.2-67.0: people are beginning to take action but may still lack confidence and skills to support new behaviors. • PAM score ≥67.1: people have confidence and perform adequate behaviors but may not be able to maintain them in the face of stress.^h
App usability						
SUS ⁱ		✓	✓	✓	For each of the 10 questions scored on a 5-point Likert scale, raw scores were obtained as follows: for odd-numbered questions, 1 was subtracted from the response value; for even-numbered questions, the response value was subtracted from 5; raw scores were converted to percentile ranks to map the raw SUS results to values calibrated from 446 studies, including >5000 individual SUS responses.	SUS score above 51 is interpreted as <i>okay</i> , with low marginal acceptability ranges; a SUS score >72 is considered acceptable, with <i>good</i> usability levels; and a SUS score >85 corresponds to <i>excellent</i> usability levels. ^j
App quality						
MARS ^k		✓	✓		MARS provides an overall mean score (each question yields a score from 1 to 5) for different dimensions (quality, functionality, esthetics, and information) of a mobile app. Only specific sections were included; section E contains 6 questions scored on a 5-point Likert scale, whereas section F contains 4 questions scored on varying scales.	No official scoring mechanism used; reported response frequencies
ACDC ^l		✓			A total of 14 questions are scored on a Likert scale from 1 to 3; only a subset of the relevant questions was extracted from the more comprehensive ACDC survey.	No official scoring mechanism used; reported response frequencies

^aHCP: health care provider.

^bEQ-5D-5L: European Quality-of-Life 5-Dimension 5-Level Questionnaire.

^cQuantity evaluated.

^dHRQoL: health-related quality of life.

^eBilbao et al [29].

^fMID: minimal important difference.

^gPAM-10: 10-item Patient Activation Measure.

^hTitova et al [35].

ⁱSUS: System Usability Scale.

^jBangor et al [36].

^kMARS: Mobile App Rating Scale.

^lACDC: App Chronic Disease Checklist.

Data Analysis

Descriptive statistics were computed for instruments measured only once at the baseline and follow-up assessments, including demographic characteristics, app usability (SUS), and app quality (MARS and ACDC), reporting the mean and SD or frequency and percentage where appropriate. Because of the small sample size, normality assumptions were not met; thus, nonparametric methods were used for all subsequent analyses. Wilcoxon matched-pairs signed-rank tests were performed to compare outcomes measured at both baseline and follow-up (PAM and EQ-5D-5L), including *P* values. In addition, app user metadata collected during the 6-week app usability period were exported, including the frequency of goal setting and activity completion. Spearman rank correlations (r_s) were used to evaluate correlations between patient activation (as measured by the PAM), number of goals set, and number of activities completed for knee OA management, with correlations <0.20 considered negligible [37].

For the descriptive thematic analysis, PaCER researchers generated descriptive statements for each app feature that

reflected the scope of the responses. Key messages were then extracted and summarized using inductive coding [38] (Multimedia Appendix 3). The notes taken during the HCP interviews were organized according to participants, questions, and app features. Using the audio recordings, the notes were then expanded upon and validated to ensure that they accurately represented the perspectives of participating HCPs. Once all interviews were completed, primary care researchers from ACTT compared similarities and differences in HCP responses and summarized the findings into descriptive categories by each app feature.

Results

Patient Evaluation

Qualitative Component

The main messages of the qualitative report generated by the research team members from PaCER (JM and ST) are summarized in Table 3. The corresponding detailed summary of the descriptive statements can be found in Multimedia Appendix 3.

Table 3. Summary of patient feedback from qualitative assessment at baseline (N=18).

App feature and areas of high likeability and usability	Areas lacking and suggestions for improvement	Implications for the patient–HCP ^a visit
Dashboard		
<ul style="list-style-type: none"> Provides a clear overall picture of the patient's knee OA^b Creates visual prompts around issues that patients may want to discuss with their HCP, such as pain in relation to activity levels Likely to improve knowledge of OA—allows patients to moderate symptoms and identify limits 	<ul style="list-style-type: none"> Should provide option for adding notes to symptom inputs Include reminders prompting patients to enter their symptoms and ability to enter data retrospectively Text should be enlarged for easier reading 	<ul style="list-style-type: none"> Considered the feature that would be most likely to improve communication with HCPs—prompts patients to discuss issues such as pain and impact on activity levels Receptiveness of HCPs was considered a limiting factor
Goals		
<ul style="list-style-type: none"> Intuitive and simple to use Useful for encouraging patients to set goals (particularly for those who set goals infrequently); helpful for self-management 	<ul style="list-style-type: none"> Goals were considered likely to work for activity and exercise, less so for symptom management Addition of reminders and a built-in reward system would be helpful Should include a notes feature so that patients can enter specifics on what was done to achieve goals Should link pain reduction goals to resources (less of a goal than an outcome) 	<ul style="list-style-type: none"> Not considered a useful feature as patients do not typically discuss goals with their HCP
Activities		
<ul style="list-style-type: none"> Easy to update, plot, and review activities; liked the calendar view Helpful to remind patients to complete activities 	<ul style="list-style-type: none"> Feature needs more specificity; should expand categories, for example, add duration to aerobic activities Present information as bullets and enlarge the font size More detailed instructions on how to use the activity feature Some overlap between activity categories (eg, aquatic and aerobic exercise) Link to exercise resources 	<ul style="list-style-type: none"> Not identified as a useful feature for communication with HCPs
Flags		
<ul style="list-style-type: none"> Provides important visual link to activity avoidance Encourages patients to be more diligent and in tune with their symptoms Helpful for patients to look back on past flags and observe changes over time Helpful for self-management and avoiding or documenting acute episodes 	<ul style="list-style-type: none"> Improve specificity of flags—expand the list of categories so they are more specific Add ≥1 descriptor per flag for improved specificity (eg, pain, low mood, and activity avoidance) 	<ul style="list-style-type: none"> Helpful to note what patients would like to discuss with their HCPs Patients expressed that their HCPs may not be keen on using this feature Logging flags would add validity to the issues patients bring to their HCPs
Information		
<ul style="list-style-type: none"> Considered the best feature by most patients—particularly the exercise resources Good information from reliable resources 	<ul style="list-style-type: none"> Include SMART^c goals link on the main page Include more strengthening exercises and guidance on exercise for specific patients (eg, with or without mobility issues) 	<ul style="list-style-type: none"> Could help to encourage discussion of local resources and information that are relevant to specific issues

^aHCP: health care provider.

^bOA: osteoarthritis.

^cSMART: specific, measurable, attainable, relevant, time-based.

In terms of overall impressions, participants responded positively to the app, noting that it was simple to use and provided a complete picture of how things were going with their OA, along with a visual record to help them keep track of their

self-management activities. Patients thought the app would motivate them to create goals, which in turn would encourage them to complete their activities. In addition, patients were keen on the *resources* feature, which included recommended exercises

and other self-management information. Patients identified the dashboard as the most effective feature for improving communication with their HCPs, providing an overall picture of their knee OA and highlighting issues to discuss with their HCPs, such as pain in relation to activity levels.

However, participants expressed some hesitancy around the potential receptiveness of HCPs to the app, suggesting that this may limit the communication potential of the app. There was wide agreement among patients that the app would likely improve their knowledge of OA, enabling them to pay more attention to their symptoms and identify their limits. Patients emphasized that the *resource* feature would be most helpful and most frequently used, although those who reported having a greater understanding of OA were less confident that it would advance their knowledge of OA. Finally, most patients were hopeful that the app would help them self-manage their knee OA, and they were tentative to make predictions until they had used the app.

In the last phase of the qualitative evaluation, patients provided specific input on how a subsequent iteration of the app may be improved. Patients suggested that simplifying the data input function, making the app more personalized to individual goals,

and enhancing the depth of content would be valuable enhancements (Table 3). Suggestions about making the app more user-friendly included more detailed instructions for each feature (including clickable information icons), expanding data input options (eg, allowing the addition of ≥ 1 activity or red flag per entry, a larger list of activity inputs, and retrospective data entry), and ensuring consistent terminology.

Recommendations from patients for improving the personalization of the app included the ability to add personal notes for activities and red flags. Feedback on enhancing the app content included adding a function to track pain medications, building in a reminder feature for activities, including more options and guidance on exercises for knee OA, and linking to more existing OA information tools (eg, My Health Alberta).

Quantitative Component

Demographic Characteristics

A total of 18 patients participated in the pilot trial, of whom 17 (94%) provided complete data at baseline and follow-up. The mean age of study participants was 62.2 (SD 6.9) years; 61% (11/18) of participants were female, and 83% (15/18) of participants had completed postsecondary education (Table 4).

Table 4. Baseline characteristics of patient participants in the app prototype evaluation (N=18).

Baseline characteristics	Values
Demographics	
Age (years), mean (SD)	66.2 (6.9)
Sex (female), n (%)	11 (61)
Postsecondary education, n (%)	15 (83)
Knee symptoms	
Probable or definite diagnosis of osteoarthritis, n (%)	16 (89)
Experienced pain, aching, or discomfort in either knee for at least a month at any time in the past, n (%)	18 (100)
Number of days per month experienced pain, stiffness, or discomfort in either or both knees, mean (SD)	22 (12)
Experienced any of the following symptoms in right knee, n (%)	
Warmth	10 (56)
Swelling	15 (83)
Redness	2 (11)
Inflammation	14 (78)
Experienced any of the following symptoms in left knee, n (%)	
Warmth	7 (39)
Swelling	11 (61)
Redness	3 (17)
Inflammation	12 (67)
Risk factors, n (%)	
Engage in physical activity at least once a week	18 (100)
Have stopped or changed the type of physical activity because of knee pain	17 (94)
Have set a goal to improve KOA ^a symptoms	10 (56)
Management of symptoms	
Have performed exercises or activities to improve KOA symptoms	14 (78)
Have ever set a goal to improve KOA symptoms	10 (56)

^aKOA: knee osteoarthritis.

Knee OA Symptoms and Risk Factors

All participants reported experiencing pain, aching, or discomfort in at least one knee within the past 12 months and reported experiencing pain, stiffness, or discomfort in either or both knees on an average of 22 (SD 11.5) days per month. Of the patients diagnosed with OA by a physician (18/18, 100%), 88% (16/18) received their diagnosis at least 1 year before participating in the study. All participants reported engaging in physical activity at least once a week, 94% (17/18) of whom reported stopping or changing the type of physical activity because of knee pain.

Knee OA Self-management

Approximately 78% (14/18) of participants indicated that they had performed exercises or activities to improve their knee OA symptoms at any point in the past, 67% (12/18) of whom had performed pain management activities or exercises >10 times in the past month. Regarding goal-setting behavior, 50% (10/18)

of patients reported previously setting a goal to improve their KOA symptoms. Of those patients, 60% (6/10) reported setting at least one goal over the 6-week evaluation period, of whom 50% (3/6) indicated a success rate >50%.

Quality of Life and Patient Activation

Between the baseline and follow-up evaluations, patient activation, measured using the PAM, increased significantly from a mean of 80.4 (SD 9.1) to 87.9 (SD 9.7; Wilcoxon signed-rank test, $P=.01$). Patient quality of life, as measured by the EQ-5D-5L, changed from a mean of 0.77 (SD 0.13) to 0.67 (SD 0.26) between week 0 and week 6 evaluations, corresponding to a significant improvement in quality of life (Wilcoxon signed-rank test, $P=.04$; Table 5). The change in EQ-5D-5L score from baseline to follow-up exceeded the minimal important difference (MID) of 0.056 based on the Canadian population [39] but did not exceed the MID of 0.20 based on the Canadian degenerative knee disease population [40].

Table 5. Results of patient-reported outcome measures collected at baseline and follow-up evaluations by patient participants.

Outcome measure	Values, mean (SD)		P value
	Baseline (n=18)	Follow-up (n=17)	
Quality of life			
EQ-5D-5L^a			
Index ^b	0.77 (0.13)	0.67 (0.26)	.04
VAS ^{c,d}	74.72 (19.36)	76.18 (17.64)	.45
Patient activation			
PAM-10 ^{e,f}	80.4 (9.1)	87.9 (9.7)	.01

^aEQ-5D-5L: European Quality-of-Life 5-Dimension 5-Level Questionnaire.

^bLow EQ-5D-5L index scores correspond to high quality of life (scale: -0.15 to 0.95).

^cVAS: visual analog scale.

^dHigh European Quality-of-Life 5-Dimension VAS scores correspond to high quality of life (scale: 0 to 100).

^ePAM-10: 10-item Patient Activation Measure.

^fHigh PAM-10 scores correspond to high patient activation (scale: 0 to 100).

App Quality and Usability

Following the 6-week pilot evaluation, patients were provided with a follow-up questionnaire to assess app usability and quality, quality of life, and patient activation. According to the responses to the ACDC assessing app quality, 53% (9/17) of patients with complete data reported that the app facilitated appropriate navigation, 65% (11/17) indicated that it was reasonably efficient, 71% (12/17) reported that it was user-friendly, and 88% (15/17) indicated that the app was free from confusing terms or jargon. Regarding the appropriate display of data and information, 88% (15/17) of patients reported that the app produced appropriate graphs or statistics for clinical data, and 77% (13/17) indicated that the app displayed correct and relevant information regarding their chronic condition and that visual explanations of concepts were clear, logical, and correct. In addition, the ACDC identified some areas of improvement for the app, with all patients indicating that the app had none or limited tactile, visual, or sound feedback, and 53% (9/17) reporting that the app did not facilitate ease of entering information. Patient responses to the MARS indicated an overall average rating of 3.1 out of 5 stars (rated on a scale of worst to best app ever used). As measured by the SUS, the mean score for perceived app usability was 57.8 (scale 0-100), indicating marginal usability according to SUS cutoffs [41].

Patient Metadata

During the 6-week observation period, user metadata were collected from the 18 participating patients, including symptoms of pain, stiffness and functional impairment, number of goals set, and number of activities completed toward knee OA

self-management. The mean number of goals set during the usability period was 2.47 (SD 3.08), and the median was 2 (IQR 1.0-3.0). The mean and median number of activities completed for knee OA self-management during the study period were 22.2 (SD 17.8) and 18.0 (IQR 5.0-41.0), respectively. Spearman rank correlation (r_s) calculations demonstrated that follow-up PAM scores were weakly correlated ($r_s=-0.32$) with the number of goals achieved and the number ($r_s=0.19$) of activities performed during the 6-week usability period.

HCP Evaluation

Qualitative Component

A detailed summary of perceived likeability and usability, suggestions for improvement, and implications for HCP visits are summarized in Table 6 (see the detailed ACTT report in Multimedia Appendix 4). Overall, HCPs expressed support for the app features and functionalities, favoring the *Goals*, *Activities*, and *Resources* tabs. HCPs identified elements of each main app feature that would be useful for patient self-management and that could be useful in the context of a clinical visit. However, HCPs largely held the perspective that the app was too detailed and cumbersome for the patient population, typically comprising older individuals perceived as having limited technological literacy. Suggestions made by the physicians to make the app more user-friendly to this patient population included increasing the use of lay language to reduce high-level language and the reading level (grade 7 or lower; 3/7, 43%), increasing color contrast and using a colorblind-friendly palette (3/7, 43%), and providing areas to add notes or free text (3/7, 43%).

Table 6. Summary of health care provider (HCP) feedback from qualitative assessment (N=7).

App feature and areas of high likeability and usability	Suggestions for improvement	Implications for the patient-HCP visit
Dashboard		
<ul style="list-style-type: none"> Helpful summary of symptoms and tracking of goals or activities Visual presentation of graphs and ability to track the completion of activities 	<ul style="list-style-type: none"> The features of different events (ie, red flags and activities) could not be identified directly from the dashboard Add additional visual features for ease of reading for patients (eg, add a legend, increase contrast and font size, and reduce reading level) 	<ul style="list-style-type: none"> Potentially too detailed to discuss within the scope of a patient visit—should be more usable at a glance Might consider highlighting pain as a main source of discussion during the patient visit (stiffness and functional impairment are less relevant)
Goals		
<ul style="list-style-type: none"> Use of the SMART^a goal-setting framework Incorporation of an assessment of confidence in achieving goals (ie, “how confident are you that you will be able to complete this goal?”)—marked on a 5-point Likert scale from not confident (1) to very confident (5) Summary of goals and prompts for next scheduled activity is useful 	<ul style="list-style-type: none"> There could be more clarity on how to use the feature—HCPs thought it might be too complex for patients to follow Improve visuals for easier reading—increase font size and color contrast Too many categories of goals Achieved goals should be removed Keep the page to one screen so that scrolling is not required 	<ul style="list-style-type: none"> Feature is most relevant for self-management
Activities		
User-friendly and straightforward data entry	<ul style="list-style-type: none"> Modify activity categories from drop-down list—make it more relatable for those who are less exercise-oriented Provide definitions for activity categories List of activity categories is limited, and language is too high level Improve readability using color contrast and different coloring 	<ul style="list-style-type: none"> Need to include an option to go back to completed activities to discuss with HCP Would be useful to see a percentage of activities completed
Flags		
Helpful for capturing activity avoidance	<ul style="list-style-type: none"> Suggestions for updates or modifications to the red flags list were provided Add option to provide notes to accompany a red flag Add option to highlight red flags to be discussed with HCP 	<ul style="list-style-type: none"> Add option to highlight red flags intended to be discussed with an HCP
Information		
<ul style="list-style-type: none"> Useful feature; considered the best tab by most physicians Information on exercise therapies was simple to understand with appropriate images and videos Reference to evidence-based programs (eg, GLAD Canada) and no equipment requirement Printable format ideal for older patients 	<ul style="list-style-type: none"> Provide exercise adaptations for patients who may be mobility-limited Provide more local resources and guidelines Add a frequently asked questions section Separate information page for resources for patients and resources for HCPs 	<ul style="list-style-type: none"> Feature is most relevant for self-management

^aSMART: specific, measurable, attainable, relevant, time-based.

Quantitative Component

App Quality and Usability

During a 1-time app evaluation, HCPs (n=7) completed a questionnaire assessing app usability and quality from the patient perspective. According to the responses to the MARS, 71% (5/7) of respondents agreed that the app was likely to increase knowledge of knee OA, 86% (6/7) of respondents indicated that

the app was likely to increase motivation to address knee OA, and 71% (5/7) of respondents considered that the app was likely to improve self-management practices. When asked if they would recommend the app to patients with OA who might benefit from it, 57% (4/7) of HCPs responded affirmatively. Overall, HCPs rated the app as 3.0 on a scale of 1.0-5.0, indicating the potential for improvement. As measured by the

SUS, the mean score for perceived app usability was 39.1, pertaining to unacceptable usability according to SUS cutoffs.

Discussion

Principal Findings

We have demonstrated that an app prototype co-designed with end users has the potential to successfully deliver self-management guidance to patients with knee OA. During the 6-week observation period, patients experienced significant improvements in patient-reported quality of life and patient activation and exhibited high levels of engagement with the app, as demonstrated by a high number of activities completed and goals achieved during the usability period. However, usability scores were reported to be in the marginal range for patients and an unacceptable range for HCPs. The overall average app ratings from patients and HCPs demonstrated the potential opportunity for further quality improvement.

The qualitative assessment revealed that, from the patient perspective, the app features, including the dashboard, goals, activities, flags, and resources, were useful and user-friendly but could be expanded to include functions that are more personalized and specific to each patient's lived experiences, such as the ability to add notes and reminders. Patients indicated that they were most likely to use the *resources* feature, emphasizing that the knowledge aspect of the tool was important to their self-management. Furthermore, patients largely viewed the app as a catalyst to increased autonomy in self-management. Many elements of the app features were viewed as useful by HCPs, particularly those related to the inclusion of evidence-based information; specific, measurable, attainable, relevant, time-based goal setting, and exercise therapies for OA.

However, there were gaps between patients' and HCPs' perceptions of app usability and quality, with HCPs expressing greater concern than patients about the patient's ability to effectively navigate and use the app features. This is consistent with our previous work [22], where physicians and patients expressed different views on the seriousness of knee OA and their approach to its management and also their perceptions of whether those diagnosed with knee OA can manage their disease using an app or other mHealth tools. Despite patient receptiveness to the prospect of using an mHealth tool such as an app, patients and physicians held diverging views, where physicians were concerned about the technological literacy of the conventionally older OA population. Negotiating consensus during co-design about the app features beneficial to both patients and physicians may help align patient and physician perspectives. Furthermore, improved patient-HCP communication and the discussion of priorities and best-available evidence may help bolster shared decision-making [42].

Limitations

We recognize several limitations of our study design. The evaluation sample size was small but aligned with the suggestion by Nielsen et al [43] that for projects of medium to large size, 15-20 test users are optimal to balance evaluation costs with the benefits of finding usability problems in testing.

Furthermore, our patients were predominantly female and had higher educational attainment. Thus, our patient sample may not be representative of the target early knee OA population, as research has shown that active participants research may have different motivations and priorities compared with those who are less engaged [44], and departures from representativeness may be amplified with increasing age [45]. This may be reflected by our study findings, in which 88% (15/17) and 100% (17/17) of participants reported PAM scores in the highest range of patient activation at baseline and follow-up assessments, respectively, despite the use of multiple recruitment mechanisms intended to capture a range of patient perspectives.

Similarly, the HCPs involved in the design and evaluation of the mHealth app were primarily family physicians. Thus, our findings may not reflect the diversity of perspectives among HCPs that are typically involved in guiding self-management for patients with knee OA. Further studies will expand the focus of the evaluation to assess its applicability in a more diverse group of HCPs (eg, physiotherapists). In addition, the inclusion of patient partners with a wider range of experiences related to education, health literacy, and technological proficiency in future app development and evaluation is essential to address potential health inequities [46].

There were shortcomings in app usability, as demonstrated by the low SUS scores reported by both patients and HCPs. However, usability challenges were of greater concern among HCPs in comparison with patients who reported higher usability scores and identified minimal to no usability challenges in the qualitative assessment. This discrepancy between patient and HCP perceptions of usability is aligned with the findings from our preliminary work [22], where HCPs largely underestimated patients' receptiveness and ability to use mHealth technologies for OA self-management. It is possible that HCPs evaluated the app usability with a broader OA patient population in mind than that represented in our sample, who were both highly educated and demonstrated high levels of activation in their care.

This study had a relatively short duration of follow-up of 6 weeks. As such, the emphasis on the evaluation of patient-reported outcomes against clinical outcomes was suitable, as it was unlikely that we would observe clinically meaningful changes in OA symptoms during the short observation period. However, although significant improvements in health-related quality of life and patient activation were reported during the study period, the change in EQ-5D-5L scores did not surpass the MID established in the degenerative knee population. Thus, the length of follow-up may have been insufficient to establish meaningful changes in some patient-reported outcomes. Given the chronic nature of OA, using a longer evaluation period for a future iteration of the app would provide more relevant data on clinical outcomes to augment the patient-reported outcomes featured here and additional context for whether long-term engagement with the app could be sustained. Further app development involving an evaluation that is more inclusive of diverse patient and HCP perspectives, with a longer follow-up duration and a wider range of outcomes measured, will be an essential next step in this study. In addition, we are exploring opportunities to integrate self-monitoring data with advanced machine learning analytics

to provide an intelligent platform to assist patients with knee OA in initiating and sustaining self-management activities.

Comparison With Previous Studies

To our knowledge, this is the first mHealth app developed specifically to aid in knee OA self-management, although several digital self-management programs have been developed for OA [47] and other chronic diseases, including diabetes [48,49] and heart failure [50]. A 2020 systematic review on digital health interventions for people with OA identified that most (5/8, 63%) studies were primarily focused on health education (n=5), whereas some incorporated additional self-management elements, such as goal setting (n=6), action planning (n=4), physical activity (n=6), weight management (n=5), and pain management (n=6) [47]. A 2017 systematic review focusing specifically on mHealth technologies identified a lack of emphasis on tracking OA symptoms and self-management behavior that could be useful for shared decision-making [9]. The review by Choi et al [9] provided a framework for developing mHealth apps for OA management, describing the need for patient-facing mobile apps with capabilities such as symptom monitoring, activity monitoring, joint function measurement, physical activity guidelines, educational content, and provision of data visualization and summary reports for shared-decision-making. Thus, our app addressed many of the identified market gaps for apps in the OA self-management domain by incorporating capabilities that have been established as integral to facilitating self-management, decision support, and shared decision-making.

Regarding methods for evaluating mHealth tools, a recent scoping review summarizing quality assessment methods for mobile apps in chronic disease management found minimal agreement on the most appropriate criteria for evaluating mobile apps, with only 18% (12/65) of apps including evidence-based health information, 22% (14/65) using a behavioral change framework, and 37% (24/65) applying usability metrics as

quality criteria for app assessment [51]. On the basis of the gaps identified in the analysis and practice methods for evaluating mobile apps, the authors proposed 3 primary goals for building quality criteria for app assessment: (1) prioritize existing evidence and knowledge against ease of assessment (eg, using patient-reported outcomes vs app ratings), (2) emphasize principles of behavior change theory, and (3) explicitly incorporate the patient perspective. Our study begins to address these identified gaps in quality assessment methods using a range of evidence-based patient-reported outcomes, incorporating behavioral change principles such as goal setting and activity monitoring, and integrating the patient perspective from early planning to app assessment phases.

Conclusions

This pilot study provides support for the use of evidence-based, patient-centered mHealth apps to improve patient-reported outcomes by encouraging patient self-efficacy and improving patient-HCP communication, ultimately promoting conservative management of knee OA. These findings, along with findings from previous study phases, provide a framework from which app developers and researchers can co-design and evaluate mHealth apps targeting self-management in a way that is inclusive of all stakeholders and reflects diverse user perspectives. Further development of the app to address usability and feasibility in the context of a larger evaluation trial, including a more diverse and representative population sample and a longer period of evaluation, will be instrumental in understanding the impact of the app on self-management and a broader range of clinical outcomes such as pain and disability. In addition, our findings can provide a basis for developing best-practice reporting standards and practices to evaluate evidence-based mHealth technologies that target self-management of chronic conditions. Future app development may involve the integration of machine learning to provide personalized self-management recommendations to patients with knee OA to address individual needs and priorities.

Acknowledgments

This study was funded by the Canadian Institutes for Health Research Knowledge to Action grant KAL147547. The funding agency was not involved in any aspect of project design or the collection, analysis, or interpretation of the data and its dissemination through the development of this manuscript.

The authors would like to acknowledge the Accelerating Change Transformation Team for their assistance with physician recruitment. The authors acknowledge the collaborative insights from Chris Smith (Alberta Bone and Joint Health Institute) and Tactica Interactive and the administrative assistance of Ms Cassandra McLaughlin for project team support.

DAM is funded through a Canada Research Chair (2008-2018) and the Arthur JE Child Chair in Rheumatology Research. KJM is a doctoral candidate in the Department of Community Health Sciences, University of Calgary, supported by the Integrated Knowledge Translation Research network, Canadian Institutes of Health Research Foundation grant (FDN #143237). Enhancing Alberta Primary Care Research Networks is funded by the Alberta Innovates Translational Health Chair Award (grant number 201300234).

Authors' Contributions

DAM, BS, LG, NM, ST, LL, TW, DCS, KM, and TB were involved in the conception and design of the study. BS, JM, ST, LG, TB, and DAM were involved in the analysis and interpretation of the results. The paper was drafted by BS and critically revised by all the authors. All authors gave final approval and agreed to be a guarantor.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Pilot testing interview guide for patients.

[[DOCX File, 16 KB - formative_v5i11e30495_app1.docx](#)]

Multimedia Appendix 2

Pilot testing interview guide for physicians.

[[DOCX File, 22 KB - formative_v5i11e30495_app2.docx](#)]

Multimedia Appendix 3

Descriptive statements from the patient qualitative evaluation (n=18).

[[DOCX File, 19 KB - formative_v5i11e30495_app3.docx](#)]

Multimedia Appendix 4

Accelerating change transformation team health care provider evaluation summary (n=7).

[[DOCX File, 20 KB - formative_v5i11e30495_app4.docx](#)]

References

1. Neogi T. The epidemiology and impact of pain in osteoarthritis. *Osteoarthritis Cartilage* 2013 Sep;21(9):1145-1153 [[FREE Full text](#)] [doi: [10.1016/j.joca.2013.03.018](https://doi.org/10.1016/j.joca.2013.03.018)] [Medline: [23973124](https://pubmed.ncbi.nlm.nih.gov/23973124/)]
2. Mora JC, Przkora R, Cruz-Almeida Y. Knee osteoarthritis: pathophysiology and current treatment modalities. *J Pain Res* 2018 Oct 5;11:2189-2196 [[FREE Full text](#)] [doi: [10.2147/JPR.S154002](https://doi.org/10.2147/JPR.S154002)] [Medline: [30323653](https://pubmed.ncbi.nlm.nih.gov/30323653/)]
3. Losina E, Daigle ME, Suter LG, Hunter DJ, Solomon DH, Walensky RP, et al. Disease-modifying drugs for knee osteoarthritis: can they be cost-effective? *Osteoarthritis Cartilage* 2013 May;21(5):655-667 [[FREE Full text](#)] [doi: [10.1016/j.joca.2013.01.016](https://doi.org/10.1016/j.joca.2013.01.016)] [Medline: [23380251](https://pubmed.ncbi.nlm.nih.gov/23380251/)]
4. Thorstenson CA, Garellick G, Rystedt H, Dahlberg LE. Better management of patients with osteoarthritis: development and nationwide implementation of an evidence-based supported osteoarthritis self-management programme. *Musculoskeletal Care* 2015 Jun;13(2):67-75. [doi: [10.1002/msc.1085](https://doi.org/10.1002/msc.1085)] [Medline: [25345913](https://pubmed.ncbi.nlm.nih.gov/25345913/)]
5. Ollivier JE. Advances in the management of osteoarthritis and rheumatoid arthritis. *JAAPA* 2001 Oct;14(10):22-5, 29-30, 36-8 (29-30, 36-8). [Medline: [11715673](https://pubmed.ncbi.nlm.nih.gov/11715673/)]
6. Devos-Comby L, Cronan T, Roesch SC. Do exercise and self-management interventions benefit patients with osteoarthritis of the knee? A metaanalytic review. *J Rheumatol* 2006 Apr;33(4):744-756. [Medline: [16583478](https://pubmed.ncbi.nlm.nih.gov/16583478/)]
7. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis Cartilage* 2007 Sep;15(9):981-1000 [[FREE Full text](#)] [doi: [10.1016/j.joca.2007.06.014](https://doi.org/10.1016/j.joca.2007.06.014)] [Medline: [17719803](https://pubmed.ncbi.nlm.nih.gov/17719803/)]
8. Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthritis Cartilage* 2019 Nov;27(11):1578-1589 [[FREE Full text](#)] [doi: [10.1016/j.joca.2019.06.011](https://doi.org/10.1016/j.joca.2019.06.011)] [Medline: [31278997](https://pubmed.ncbi.nlm.nih.gov/31278997/)]
9. Choi W, Zheng H, Franklin P, Tulu B. mHealth technologies for osteoarthritis self-management and treatment: a systematic review. *Health Informatics J* 2019 Sep;25(3):984-1003 [[FREE Full text](#)] [doi: [10.1177/1460458217735676](https://doi.org/10.1177/1460458217735676)] [Medline: [29090628](https://pubmed.ncbi.nlm.nih.gov/29090628/)]
10. Lenzen SA, Daniëls R, van Bokhoven MA, van der Weijden T, Beurskens A. Disentangling self-management goal setting and action planning: a scoping review. *PLoS One* 2017 Nov 27;12(11):e0188822 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0188822](https://doi.org/10.1371/journal.pone.0188822)] [Medline: [29176800](https://pubmed.ncbi.nlm.nih.gov/29176800/)]
11. 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)] [Medline: [12867348](https://pubmed.ncbi.nlm.nih.gov/12867348/)]
12. Lorig K. Action planning: a call to action. *J Am Board Fam Med* 2006;19(3):324-325 [[FREE Full text](#)] [doi: [10.3122/jabfm.19.3.324](https://doi.org/10.3122/jabfm.19.3.324)] [Medline: [16672687](https://pubmed.ncbi.nlm.nih.gov/16672687/)]
13. Gay C, Chabaud A, Guilley E, Coudeyre E. Educating patients about the benefits of physical activity and exercise for their hip and knee osteoarthritis. Systematic literature review. *Ann Phys Rehabil Med* 2016 Jun;59(3):174-183 [[FREE Full text](#)] [doi: [10.1016/j.rehab.2016.02.005](https://doi.org/10.1016/j.rehab.2016.02.005)] [Medline: [27053003](https://pubmed.ncbi.nlm.nih.gov/27053003/)]
14. Dahlberg LE, Grahn D, Dahlberg JE, Thorstenson CA. A web-based platform for patients with osteoarthritis of the hip and knee: a pilot study. *JMIR Res Protoc* 2016 Jun 03;5(2):e115 [[FREE Full text](#)] [doi: [10.2196/resprot.5665](https://doi.org/10.2196/resprot.5665)] [Medline: [27261271](https://pubmed.ncbi.nlm.nih.gov/27261271/)]

15. Pelle T, Bevers K, van der Palen J, van den Hoogen FH, van den Ende CH. Development and evaluation of a tailored e-self-management intervention (dr. Bart app) for knee and/or hip osteoarthritis: study protocol. *BMC Musculoskelet Disord* 2019 Aug 31;20(1):398 [FREE Full text] [doi: [10.1186/s12891-019-2768-9](https://doi.org/10.1186/s12891-019-2768-9)] [Medline: [31472687](https://pubmed.ncbi.nlm.nih.gov/31472687/)]
16. Joosten E, DeFuentes-Merillas L, de Weert GH, Sensky T, van der Staak CP, de Jong CA. Systematic review of the effects of shared decision-making on patient satisfaction, treatment adherence and health status. *Psychother Psychosom* 2008;77(4):219-226 [FREE Full text] [doi: [10.1159/000126073](https://doi.org/10.1159/000126073)] [Medline: [18418028](https://pubmed.ncbi.nlm.nih.gov/18418028/)]
17. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. *PLoS One* 2016 May 23;11(5):e0156164 [FREE Full text] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]
18. Subhi Y, Bube SH, Rolskov Bojsen S, Skou Thomsen AS, Konge L. Expert involvement and adherence to medical evidence in medical mobile phone apps: a systematic review. *JMIR Mhealth Uhealth* 2015 Jul 27;3(3):e79 [FREE Full text] [doi: [10.2196/mhealth.4169](https://doi.org/10.2196/mhealth.4169)] [Medline: [26215371](https://pubmed.ncbi.nlm.nih.gov/26215371/)]
19. 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/)]
20. Farago P. App engagement: the matrix reloaded. Flurry. 2012. URL: <https://www.flurry.com/blog/app-engagement-the-matrix-reloaded/> [accessed 2021-06-02]
21. Anderson K, Emmerton LM. Contribution of mobile health applications to self-management by consumers: review of published evidence. *Aust Health Rev* 2016 Nov;40(5):591-597. [doi: [10.1071/AH15162](https://doi.org/10.1071/AH15162)] [Medline: [26681206](https://pubmed.ncbi.nlm.nih.gov/26681206/)]
22. Barber T, Sharif B, Teare S, Miller J, Shewchuk B, Green LA, et al. Qualitative study to elicit patients' and primary care physicians' perspectives on the use of a self-management mobile health application for knee osteoarthritis. *BMJ Open* 2019 Feb 01;9(1):e024016 [FREE Full text] [doi: [10.1136/bmjopen-2018-024016](https://doi.org/10.1136/bmjopen-2018-024016)] [Medline: [30782723](https://pubmed.ncbi.nlm.nih.gov/30782723/)]
23. Mrklas KJ, Barber T, Campbell-Scherer D, Green LA, Li LC, Marlett N, et al. Co-design in the development of a mobile health app for the management of knee osteoarthritis by patients and physicians: qualitative study. *JMIR Mhealth Uhealth* 2020 Jul 10;8(7):e17893 [FREE Full text] [doi: [10.2196/17893](https://doi.org/10.2196/17893)] [Medline: [32673245](https://pubmed.ncbi.nlm.nih.gov/32673245/)]
24. 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/)]
25. Given LM. *The SAGE Encyclopedia of Qualitative Research Methods*. Los Angeles: California Publications; 2008.
26. Rogers EM. *Diffusion of Innovations*. 5th Edition. New York: Free Press; 2003.
27. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. *Int J Med Inform* 2004 Nov;73(11-12):781-795 [FREE Full text] [doi: [10.1016/j.ijmedinf.2004.08.003](https://doi.org/10.1016/j.ijmedinf.2004.08.003)] [Medline: [15491929](https://pubmed.ncbi.nlm.nih.gov/15491929/)]
28. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. *IEEE Trans Profess Commun* 2000;43(3):261-278 [FREE Full text] [doi: [10.1109/47.867942](https://doi.org/10.1109/47.867942)]
29. Bilbao A, García-Pérez L, Arenaza JC, García I, Ariza-Cardiel G, Trujillo-Martín E, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual Life Res* 2018 Nov;27(11):2897-2908 [FREE Full text] [doi: [10.1007/s11136-018-1929-x](https://doi.org/10.1007/s11136-018-1929-x)] [Medline: [29978346](https://pubmed.ncbi.nlm.nih.gov/29978346/)]
30. The Patient Activation Measure® (PAM®). Insignia Health. 2020. URL: <https://s3.amazonaws.com/insigniahealth.com-assets/PAM-Fact-Sheet.20200505.pdf?mtime=20200505094829&focal=none> [accessed 2021-07-12]
31. Hibbard J. Assessing measurement properties of the PAM-10 and PAM-13. University of Oregon. 2017. URL: <https://s3.amazonaws.com/insigniahealth.com-assets/Comparing-PAM-10-to-PAM-13.pdf?mtime=20191206162413> [accessed 2021-03-10]
32. Vlachogianni P, Tselios N. Perceived usability evaluation of educational technology using the System Usability Scale (SUS): a systematic review. *J Res Technol Educ* 2021 Feb 23:1-18 [FREE Full text] [doi: [10.1080/15391523.2020.1867938](https://doi.org/10.1080/15391523.2020.1867938)]
33. Anderson K, Burford O, Emmerton L. App chronic disease checklist: protocol to evaluate mobile apps for chronic disease self-management. *JMIR Res Protoc* 2016 Nov 04;5(4):e204 [FREE Full text] [doi: [10.2196/resprot.6194](https://doi.org/10.2196/resprot.6194)] [Medline: [27815233](https://pubmed.ncbi.nlm.nih.gov/27815233/)]
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. Titova E, Salvesen Ø, Bentsen SB, Sunde S, Steinshamn S, Henriksen AH. Does an integrated care intervention for COPD patients have long-term effects on quality of life and patient activation? A prospective, open, controlled single-center intervention study. *PLoS One* 2017 Jan 6;12(1):e0167887 [FREE Full text] [doi: [10.1371/journal.pone.0167887](https://doi.org/10.1371/journal.pone.0167887)] [Medline: [28060921](https://pubmed.ncbi.nlm.nih.gov/28060921/)]
36. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 2009;4(3):114-123 [FREE Full text]
37. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg* 2018 May;126(5):1763-1768. [doi: [10.1213/ANE.0000000000002864](https://doi.org/10.1213/ANE.0000000000002864)] [Medline: [29481436](https://pubmed.ncbi.nlm.nih.gov/29481436/)]
38. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods* 2006;5(1):80-92 [FREE Full text] [doi: [10.1177/160940690600500107](https://doi.org/10.1177/160940690600500107)]

39. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-defined estimates of the minimally important difference for EQ-5D-5L index scores. *Value Health* 2017 Apr;20(4):644-650 [FREE Full text] [doi: [10.1016/j.jval.2016.11.015](https://doi.org/10.1016/j.jval.2016.11.015)] [Medline: [28408007](https://pubmed.ncbi.nlm.nih.gov/28408007/)]
40. Devji T, Guyatt GH, Lytvyn L, Brignardello-Petersen R, Foroutan F, Sadeghirad B, et al. Application of minimal important differences in degenerative knee disease outcomes: a systematic review and case study to inform BMJ Rapid Recommendations. *BMJ Open* 2017 May 11;7(5):e015587 [FREE Full text] [doi: [10.1136/bmjopen-2016-015587](https://doi.org/10.1136/bmjopen-2016-015587)] [Medline: [28495818](https://pubmed.ncbi.nlm.nih.gov/28495818/)]
41. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012 Dec 15;380(9859):2197-2223. [doi: [10.1016/S0140-6736\(12\)61689-4](https://doi.org/10.1016/S0140-6736(12)61689-4)] [Medline: [23245608](https://pubmed.ncbi.nlm.nih.gov/23245608/)]
42. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med* 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: [10.1007/s11606-012-2077-6](https://doi.org/10.1007/s11606-012-2077-6)] [Medline: [22618581](https://pubmed.ncbi.nlm.nih.gov/22618581/)]
43. 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 Presented at: INTERCHI93: Conference on Human Factors in Computing; Apr 24 - 29, 1993; Amsterdam The Netherlands. [doi: [10.1145/169059.169166](https://doi.org/10.1145/169059.169166)]
44. Maguire K, Britten N. "How can anybody be representative for those kind of people?" Forms of patient representation in health research, and why it is always contestable. *Soc Sci Med* 2017 Jun;183:62-69 [FREE Full text] [doi: [10.1016/j.socscimed.2017.04.049](https://doi.org/10.1016/j.socscimed.2017.04.049)] [Medline: [28463721](https://pubmed.ncbi.nlm.nih.gov/28463721/)]
45. Golomb BA, Chan VT, Evans MA, Koperski S, White HL, Criqui MH. The older the better: are elderly study participants more non-representative? A cross-sectional analysis of clinical trial and observational study samples. *BMJ Open* 2012 Dec 14;2(6):e000833 [FREE Full text] [doi: [10.1136/bmjopen-2012-000833](https://doi.org/10.1136/bmjopen-2012-000833)] [Medline: [23242479](https://pubmed.ncbi.nlm.nih.gov/23242479/)]
46. Abbasgholizadeh Rahimi S, Menear M, Robitaille HF, Légaré F. Are mobile health applications useful for supporting shared decision making in diagnostic and treatment decisions? *Glob Health Action* 2017 Jun;10(3):1332259 [FREE Full text] [doi: [10.1080/16549716.2017.1332259](https://doi.org/10.1080/16549716.2017.1332259)] [Medline: [28838306](https://pubmed.ncbi.nlm.nih.gov/28838306/)]
47. Safari R, Jackson J, Sheffield D. Digital self-management interventions for people with osteoarthritis: systematic review with meta-analysis. *J Med Internet Res* 2020 Jul 20;22(7):e15365 [FREE Full text] [doi: [10.2196/15365](https://doi.org/10.2196/15365)] [Medline: [32706657](https://pubmed.ncbi.nlm.nih.gov/32706657/)]
48. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. *J Med Internet Res* 2012 May 08;14(3):e70 [FREE Full text] [doi: [10.2196/jmir.2058](https://doi.org/10.2196/jmir.2058)] [Medline: [22564332](https://pubmed.ncbi.nlm.nih.gov/22564332/)]
49. Bradway M, Morris RL, Giordanengo A, Årsand E. How mHealth can facilitate collaboration in diabetes care: qualitative analysis of co-design workshops. *BMC Health Serv Res* 2020 Nov 30;20(1):1104 [FREE Full text] [doi: [10.1186/s12913-020-05955-3](https://doi.org/10.1186/s12913-020-05955-3)] [Medline: [33256732](https://pubmed.ncbi.nlm.nih.gov/33256732/)]
50. Woods L, Duff J, Roehrer E, Walker K, Cummings E. Design of a consumer mobile health app for heart failure: findings from the nurse-led co-design of Care4myHeart. *JMIR Nurs* 2019;2(1):e14633 [FREE Full text] [doi: [10.2196/14633](https://doi.org/10.2196/14633)] [Medline: [34345774](https://pubmed.ncbi.nlm.nih.gov/34345774/)]
51. Agarwal P, Gordon D, Griffith J, Kithulegoda N, Witteman HO, Sacha Bhatia R, et al. Assessing the quality of mobile applications in chronic disease management: a scoping review. *NPJ Digit Med* 2021 Mar 10;4(1):46 [FREE Full text] [doi: [10.1038/s41746-021-00410-x](https://doi.org/10.1038/s41746-021-00410-x)] [Medline: [33692488](https://pubmed.ncbi.nlm.nih.gov/33692488/)]

Abbreviations

- ACDC:** App Chronic Disease Checklist
- ACTT:** Accelerating Change Transformation Team
- EQ-5D-5L:** European Quality-of-Life 5-Dimension 5-Level Questionnaire
- HCP:** health care provider
- MARS:** Mobile App Rating Scale
- mHealth:** mobile health
- MID:** minimal important difference
- OA:** osteoarthritis
- PaCER:** Patient and Community Engagement Research
- PAM:** patient activation measure
- SUS:** System Usability Scale

Edited by G Eysenbach; submitted 01.06.21; peer-reviewed by C Cheng, C Jacob; comments to author 09.08.21; revised version received 03.10.21; accepted 07.10.21; published 25.11.21.

Please cite as:

Shewchuk B, Green LA, Barber T, Miller J, Teare S, Campbell-Scherer D, Mrklas KJ, Li LC, Marlett N, Wasylak T, Lopatina E, McCaughey D, Marshall DA

Patients' Use of Mobile Health for Self-management of Knee Osteoarthritis: Results of a 6-Week Pilot Study

JMIR Form Res 2021;5(11):e30495

URL: <https://formative.jmir.org/2021/11/e30495>

doi: [10.2196/30495](https://doi.org/10.2196/30495)

PMID: [34842526](https://pubmed.ncbi.nlm.nih.gov/34842526/)

©Brittany Shewchuk, Lee A Green, Tanya Barber, Jean Miller, Sylvia Teare, Denise Campbell-Scherer, Kelly J Mrklas, Linda C Li, Nancy Marlett, Tracy Wasylak, Elena Lopatina, Deirdre McCaughey, Deborah A Marshall. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.11.2021. 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

Contemporary English Pain Descriptors as Detected on Social Media Using Artificial Intelligence and Emotion Analytics Algorithms: Cross-sectional Study

Ming Yi Tan¹, BDS, MDS; Charlene Enhui Goh¹, BDS, MPH, DrPH; Hee Hon Tan¹, BDS, MDS

Faculty of Dentistry, National University of Singapore, Singapore, Singapore

Corresponding Author:

Ming Yi Tan, BDS, MDS

Faculty of Dentistry, National University of Singapore

9 Lower Kent Ridge Road

Singapore, 119085

Singapore

Phone: 65 67725340

Email: dentmy@nus.edu.sg

Abstract

Background: Pain description is fundamental to health care. The McGill Pain Questionnaire (MPQ) has been validated as a tool for the multidimensional measurement of pain; however, its use relies heavily on language proficiency. Although the MPQ has remained unchanged since its inception, the English language has evolved significantly since then. The advent of the internet and social media has allowed for the generation of a staggering amount of publicly available data, allowing linguistic analysis at a scale never seen before.

Objective: The aim of this study is to use social media data to examine the relevance of pain descriptors from the existing MPQ, identify novel contemporary English descriptors for pain among users of social media, and suggest a modification for a new MPQ for future validation and testing.

Methods: All posts from social media platforms from January 1, 2019, to December 31, 2019, were extracted. Artificial intelligence and emotion analytics algorithms (Crystalace and CrystalFeel) were used to measure the emotional properties of the text, including *sarcasm*, *anger*, *fear*, *sadness*, *joy*, and *valence*. Word2Vec was used to identify new pain descriptors associated with the original descriptors from the MPQ. Analysis of count and pain intensity formed the basis for proposing new pain descriptors and determining the order of pain descriptors within each subclass.

Results: A total of 118 new associated words were found via Word2Vec. Of these 118 words, 49 (41.5%) words had a count of at least 110, which corresponded to the count of the bottom 10% (8/78) of the original MPQ pain descriptors. The count and intensity of pain descriptors were used to formulate the inclusion criteria for a new pain questionnaire. For the suggested new pain questionnaire, 11 existing pain descriptors were removed, 13 new descriptors were added to existing subclasses, and a new *Psychological* subclass comprising 9 descriptors was added.

Conclusions: This study presents a novel methodology using social media data to identify new pain descriptors and can be repeated at regular intervals to ensure the relevance of pain questionnaires. The original MPQ contains several potentially outdated pain descriptors and is inadequate for reporting the psychological aspects of pain. Further research is needed to examine the reliability and validity of the revised MPQ.

(JMIR Form Res 2021;5(11):e31366) doi:[10.2196/31366](https://doi.org/10.2196/31366)

KEYWORDS

pain descriptors; social media; artificial intelligence; emotion analytics; McGill Pain Questionnaire

Introduction

Pain is “an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage” [1]. Although the experience of pain

is common, each person’s pain is unique and felt physically only by that person. Pain measurement is essential for diagnosis, monitoring of disease progression, and evaluation of treatment effectiveness [2]. This understanding and measurement of pain occur largely through verbal reporting by the person

experiencing the pain. Health care providers use this information to characterize and determine pain intensity, and in many instances, make judgments regarding treatment [3]. As verbalization of pain can sometimes be difficult, comprehensive questionnaires to elicit better descriptions of pain for diagnostic accuracy are important.

Various instruments have been developed for the assessment of pain. For acute pain, pain scales that focus on identifying pain location and intensity, such as the visual analog scale and numeric rating scale, are most commonly used [4]. Although shown to be inferior to both the visual analog scale and numeric rating scale, 4-point or 6-point verbal categorical rating scales using adjectives to describe different levels of pain have also been successfully used [5]. For children, pain scales that use images of happy and unhappy faces have been found to be appropriate [6]. All these instruments have been validated in clinical and research settings; however, their accuracy is dependent on timely recording.

Assessment of chronic pain is indisputably more complex. The long-term burden of pain plays a profound role in shaping an individual's physical and psychological state. In addition, the negative downstream effects of chronic pain can exacerbate the original pain condition through various pathways that remain poorly understood. Owing to these complexities, the above unidimensional instruments that only describe pain in terms of intensity may be too simplistic for meaningful clinical correlation.

Several authors have emphasized the need to recognize the multidimensional aspects of pain [7-9], and 6 dimensions have previously been described: physiological, sensory, affective, cognitive, behavioral, and sociocultural [7]. Health care professionals have acknowledged the varying qualities of pain and often depend on characteristic descriptions of pain to identify eventual diagnoses.

The McGill Pain Questionnaire (MPQ), created in 1971, is one of the most frequently cited instruments and has been validated for use in asymptomatic, symptomatic, and persistently symptomatic populations [10]. Its role in clinical research has been established, in large part because of the multidimensional nature of this instrument [7,11]. The original MPQ [12,13] comprises 78 pain descriptors broadly divided into 3 major classes: sensory, affective, and evaluative. These pain descriptors are further categorized into 20 subclasses, with the words within the respective subclasses ranked based on pain intensity.

The MPQ can be administered by an interviewer who reads the instructions to the patient and defines any words that are not understood [14]. It may also be self-administered. There is heavy reliance on language proficiency, a lack of which may limit the effectiveness of the instrument. Regional linguistic patterns may also be inadvertently incorporated into practice. In addition, language itself is constantly changing and evolving, with use and meanings constantly being updated.

The MPQ has remained unchanged since its inception, and the impact of modern language use on the relevance of MPQ pain descriptors remains unreported. During this time, the invention

of the internet and its exponential penetration have drastically reshaped our world. Social media, which comprises forums, blogs, business networks, social gaming, microblogs, photo-sharing platforms, and chat apps, has evolved dramatically alongside these developments. Furthermore, >50% of the global population was expected to access the internet in 2019, and the same figure was expected to use social media platforms [15]. This generates a staggering amount of publicly available social media data, allowing linguistic analysis at a scale that has never been seen before.

Natural language processing or computational linguistics, together with machine learning algorithms, have evolved substantially over the years to be able to analyze, learn, and understand the linguistic contexts of words, identify sentiments and emotions, and form neural network models [16]. These have enabled health researchers to use social media data to address a wide range of public health concerns, including surveillance and inference of developing trends in infectious diseases, targeting of public health strategies, evaluation of public health interventions, ascertaining public perceptions of nonmedical use of opioids, and even predicting disease status [17-21].

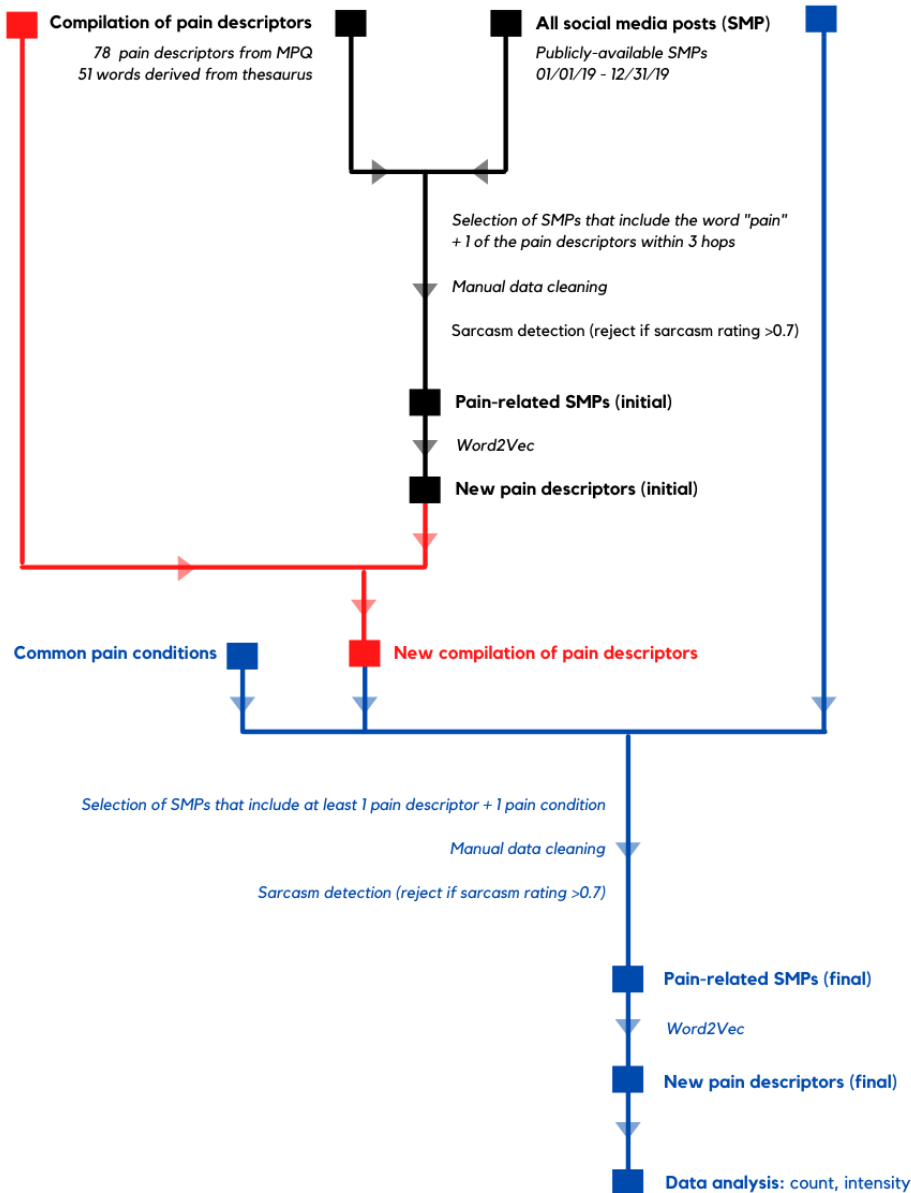
In relation to chronic pain, social media represents a snapshot of natural day-to-day colloquial language rather than formal communication. Furthermore, the nature of social media encourages users to capture their thoughts and ideas instantaneously. This is particularly important for accurate pain reporting, which is vulnerable to recall bias and relies heavily on timely reporting. Previous studies have observed extensive amounts of web-based conversations regarding pain [22], suggesting that social media may be a valid place to obtain a large amount of data regarding pain experience. Another study demonstrated that the characteristics of pain conditions could be discerned from social media posts [23], indicating that the depth of data may be sufficient for health care workers to further analyze and understand pain. Therefore, the objectives of this investigation are to use social media data to examine the relevance of pain descriptors from the existing MPQ, identify novel contemporary English descriptors for pain among users of social media, and suggest a modification for a new MPQ for future validation and testing.

Methods

Overview

This study used artificial intelligence and emotion analytics algorithms for the derivation and analysis of pain expression from social media platforms. The workflow was executed by a company specializing in linguistic and emotion analytics (INTNT.AI) and comprised 5 main steps conducted in an iterative process: (1) preliminary data gathering, (2) data cleaning, (3) Word2Vec (patent number US9037464B1; Google Inc), (4) final data gathering and cleaning, and (5) data analysis. The workflow is summarized in Figure 1. Preliminary data gathering was first performed to identify new pain descriptors through Word2Vec, which could then be used to maximize the search for relevant social media posts in the final data gathering.

Figure 1. Workflow for the derivation of new pain descriptors. SMP: social media post.



All posts from social media platforms over a 1-year period, from January 1, 2019, to December 31, 2019, were extracted. These data were acquired from a social listening platform (Meltwater) that aggregates and gives direct and official access to all accounts open to the public on Twitter, Facebook, Instagram, and YouTube.

Preliminary Data Gathering

A list comprising 78 pain descriptors from the MPQ and 51 additions yielded through the use of a mixture of web-based thesauruses was used to identify relevant social media posts (Multimedia Appendix 1). Word networks were created from social media posts using word graphs and path search algorithms that allow the linking of words to the root word of interest [24,25]. Words immediately next to the root word are described as one *hop* away, whereas subsequent layers of words are in increasing levels of hops.

Only social media posts containing the word *pain* and one of the targeted pain descriptors within 3 *hops* were included in this

study. The criterion of within 3 *hops* was decided based on pilot testing that showed good richness of social media posts that balanced relevance with the breadth of content.

Data Cleaning of the Social Media Posts

The selected social media posts were manually scrutinized and cleaned. Usernames, hyperlinks, and internet-specific symbols were removed. In addition, the content of the social media posts was evaluated for relevance. Social media posts that contained the pain descriptor in irrelevant contexts were removed.

The remaining posts were then input into a sarcasm detection machine (Crystalace, Institute of High Performance Computing, Agency for Science, Technology and Research), which is a support vector machine classifier trained with an affect-cognition-sociolinguistics feature model [26,27]. The Crystalace sarcasm detection method rated social media posts for sarcasm on a scale of 0-1, with 1 indicating maximum sarcasm. Social media posts with a sarcasm rating >0.7 were removed from the database.

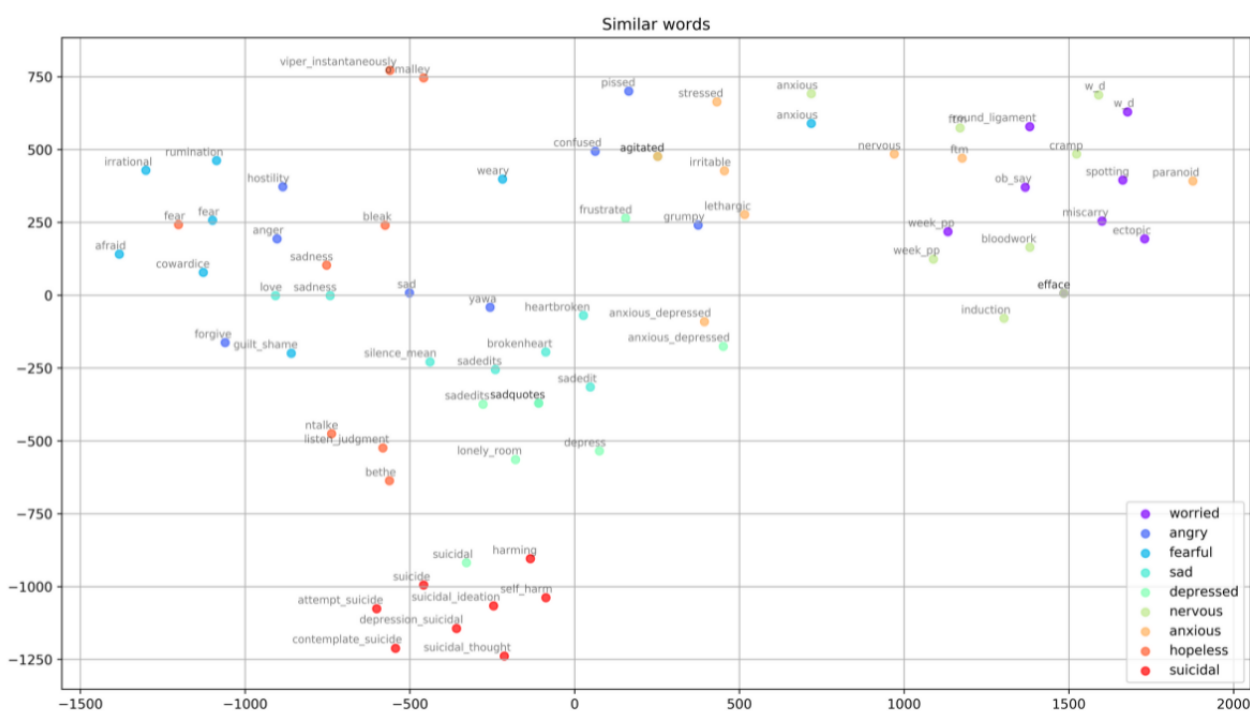
Sarcasm is a difficult concept to handle in emotion analysis. Traditionally, sarcasm has been viewed from a psychological perspective where overt irony is actively pursued by the speaker as a tool of *verbal violence* [28]. To achieve the required effect and aggression, sarcasm requires a higher level of semantic complexity than normal conversation, in which both positive and negative connotations are interwoven. The CrystalNest software (CrystalNest, Institute of High Performance Computing, Agency for Science, Technology and Research) combines feature engineering and deep learning frameworks with linguistic rules overlay. This performed very well in sarcasm detection for SEMEVAL 2017 and 2018 and was

therefore selected for this study [24,25]. Further testing of later models of the CrystalFeel emotion intensity analysis engine also indicated accuracy levels of 0.818, 0.765, 0.765, 0.788, and 0.856 for predicting the intensities of anger, fear, sadness, joy, and valence, respectively [29].

Word2Vec to Identify Associated Words

Word2Vec [30] was then used on the remaining social media posts to convert the semantic meaning of words into a numerical representation, or vectors, based on the context in which those words occur. Words that share common contexts are located close to one another within the overall vector space constructed from the inputted text (Figure 2).

Figure 2. Sample 2D illustration of words with common contexts within the overall 3D vector space; similar words are color-coded. This graph was constructed using T-distributed Stochastic Neighbor Embedding (TSNE), to aid visualization of word clusters. TSNE works by taking a group of high-dimensional vocabulary word feature vectors, then compressing them down to 2-dimensional x, y coordinate pairs. This method keeps similar words close together on the plane, while maximizing the distance between dissimilar words.



This process allowed for the classification of pain dimensions from the open text found in the included social media posts and the identification of new pain descriptors. All new words with a positive vector distance from the original list of pain descriptors were considered to be associated, and a maximum of 20 associated words per root word was selected for inclusion.

Final Data Gathering and Cleaning

Newly identified pain descriptors derived from Word2Vec mechanisms were compiled with the original list of pain descriptors used in preliminary data gathering to form an expanded list of keywords to be used for the final round of data gathering. The search for relevant social media posts was performed in the same social media platforms and period as above.

For greater specificity to health conditions, social media posts were included only if they contained at least one of the pain descriptors in this new list, as well as one pain condition from a list of common pain conditions (Multimedia Appendix 2).

This list of pain conditions was not input in the preliminary round of data gathering to ensure maximal identification of possible pain descriptors, regardless of context. The combination of pain descriptors with pain conditions eliminated most of the social media posts that were irrelevant to the study.

The final data set was obtained after the selected social media posts were put through the same data cleaning steps as detailed above.

Data Analysis

The original MPQ comprised 78 pain descriptors categorized into 20 subclasses. In addition, 51 additional words were yielded through the use of web-based thesauruses. These 129 original pain descriptors served as keywords for the final analysis. The final data set was input into Word2Vec, and the final classification of descriptors was obtained using the algorithm.

Words found to be related to keywords were identically color-coded, located in close proximity to the overall vector

space, and had a positive vector distance. A maximum of 20 words found to be most similar to each of the predetermined 129 keywords was selected for inclusion and further pruning. These were evaluated for relevance to pain descriptors. Entries with contrasting meanings to the keywords (eg, *hot vs cold*), entries that shared the same root word as the MPQ keywords, and entries irrelevant to pain description were removed.

The number of mentions, or count, of each pain descriptor within the final data set of selected social media posts was computed. Analysis of the counts for the 78 original MPQ keywords was conducted to determine the minimum threshold level for the inclusion of new associated words. The bottom 10% of the original MPQ pain descriptors were found to have counts of <110. Therefore, a count of 110 was set as the minimum threshold level for the prevalence of word use, and all words with a count of <110 were removed.

The intensities of all descriptors were also analyzed using natural language processing emotion analytics algorithms (CrystalFeel, Institute of High Performance Computing, Agency for Science, Technology and Research), which considered the entire sentence or paragraph in which each descriptor was found. The CrystalFeel algorithms allowed for the measurement of the emotional properties in text, including *anger, fear, sadness, joy, and valence*. Intensity was reflected on a scale of 0-1, with 1 implying maximum pain intensity. This formed the basis for determining the order of pain descriptors within each subclass.

Results

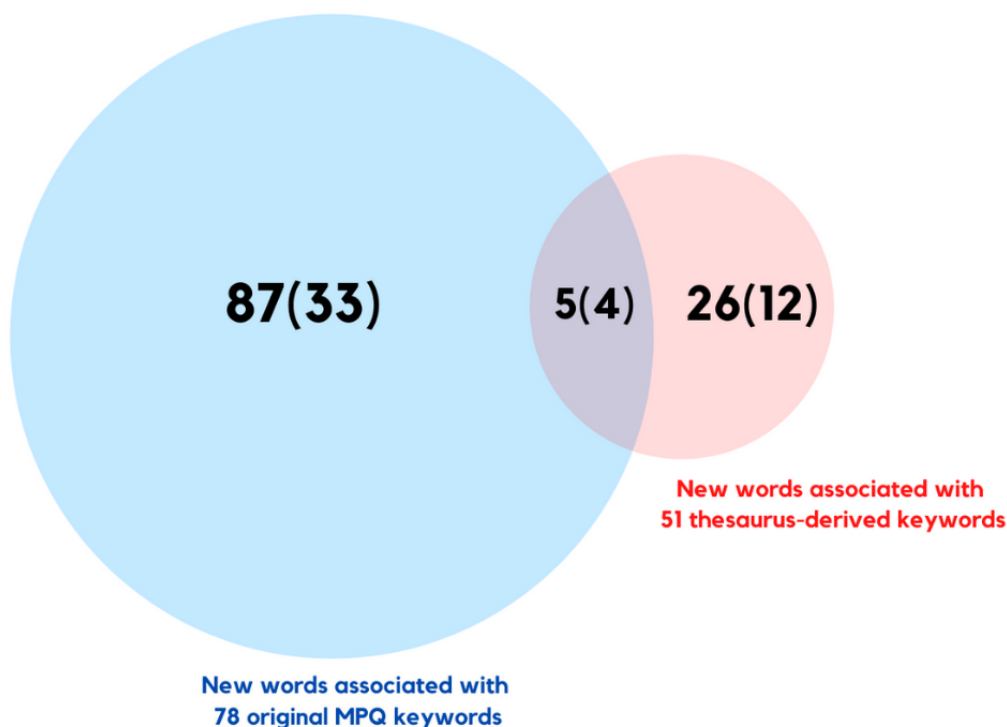
Sample of Social Media Posts

A total of 572,742 social media posts were obtained from a preliminary round of data gathering. Following manual evaluation of the 572,742 social media posts for relevance, 8310 (1.45%) social media posts were removed, whereas 7824 (1.37%) social media posts were removed after failing to meet the threshold criterion for sarcasm. Word2Vec identified 34 new pain descriptors that were used together with the original list of words to widen the search for relevant social media posts in the second round of data gathering. A total of 1,877,122 social media posts were identified in the second round of data gathering. After data cleaning and the additional inclusion criteria of containing at least one pain condition as well as one pain descriptor in the social media post, 11.55% (216,873/1,877,122) of social media posts remained for the final data analysis.

Identification of New Pain Descriptors Through Word2Vec

Using Word2Vec, a total of 118 new associated words were found for the 129 pain descriptor keywords defined in this study, following the removal of repetitions. Of these 118 words, 5 (4.2%) were associated with both the original MPQ and thesaurus-derived keywords, 87 (73.7%) were associated with MPQ keywords only, and 26 (22%) were associated with thesaurus-derived keywords only (Figure 3).

Figure 3. A total of 118 new words were found to be associated with the 78 original McGill Pain Questionnaire keywords and 51 thesaurus-derived keywords; bracketed values indicate the number of words with count ≥110. MPQ: McGill Pain Questionnaire.



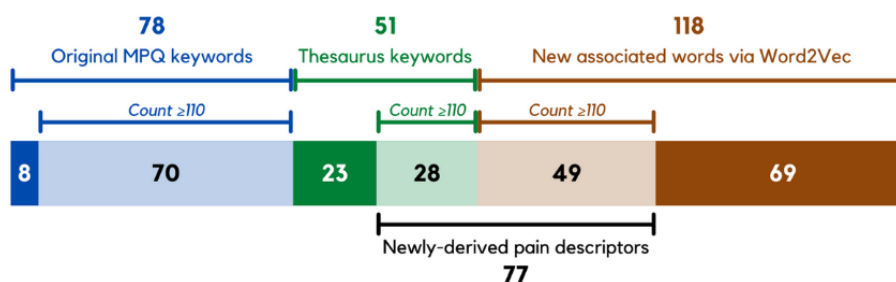
Count and Intensity of the Original MPQ Keywords and New Pain Descriptors

Of the 118 new associated words acquired through Word2Vec, 49 (41.5%) words were found to have a count of at least 110, meeting the minimum threshold level for the prevalence of word use.

28 (23.7%) thesaurus-derived keywords met the minimum threshold count of 110. These and the 49 out of 118 (41.5%)

new associated words derived through Word2Vec were combined into a single list for further evaluation, representing the final list of 77 newly derived pain descriptors (Figure 4). From this list, the pain descriptors that received the top 10 highest counts in descending order were *anxiety* (96909), *depression* (65223), *fear* (48165), *excruciating* (32094), *anger* (31506), *discomfort* (31333), *depress* (29822), *sadness* (25817), *low* (23027), and *fever* (21441).

Figure 4. Breakdown of original and newly derived pain descriptors. MPQ: McGill Pain Questionnaire.



Of the 78 original MPQ keywords, the pain descriptors that received the top 10 highest counts in descending order were *sharp* (26679), *hot* (25405), *tingling* (25405), *intense* (24269), *numb* (23964), *unbearable* (20929), *cold* (19099), *sore* (16127), *burning* (14204), and *itchy* (14204).

The intensity of the 78 original MPQ keywords ranged from 0.367 (for *flickering* and *jumping*) to 0.699 (for *terrifying*). For the combined list of 77 newly derived pain descriptors, the intensity ranged from 0.325 (for *scratching*) to 0.7467 (for *suicidal*).

The counts and intensities of existing and new words are presented in Multimedia Appendix 3.

Suggested New Pain Questionnaire

Overview

Analyses of count and intensity of the pain descriptors, as well as manual analysis of the social media posts for the context in which the word was used, were used to recommend the inclusion of pain descriptors in a new pain questionnaire. The suggested changes, categorized by the MPQ subclasses, are summarized in Table 1.

Table 1. Comparison of the original McGill Pain Questionnaire (MPQ) with the suggestion for a new pain questionnaire.

Subclass and pain descriptors from the original MPQ	Suggested words for a new pain questionnaire	Ranking reordered? (yes or no)
Temporal		
<ul style="list-style-type: none"> • Flickering • Quivering (removed) • Pulsing • Throbbing • Beating (removed) • Pounding 	<ul style="list-style-type: none"> • Flickering • Pulsing • Throbbing • Pounding 	<ul style="list-style-type: none"> • No
Spatial		
<ul style="list-style-type: none"> • Jumping • Flashing • Shooting 	<ul style="list-style-type: none"> • Jumping • Flashing • Shooting 	<ul style="list-style-type: none"> • No
Punctuate pressure		
<ul style="list-style-type: none"> • Pricking • Boring • Drilling • Stabbing • Lancing (removed) 	<ul style="list-style-type: none"> • Pricking • Boring • Drilling • Stabbing • Puncturing (new) 	<ul style="list-style-type: none"> • No
Incisive pressure		
<ul style="list-style-type: none"> • Sharp • Cutting • Lacerating (removed) 	<ul style="list-style-type: none"> • Cutting • Sharp 	<ul style="list-style-type: none"> • Yes
Constrictive pressure		
<ul style="list-style-type: none"> • Pinching • Pressing • Gnawing • Cramping • Crushing 	<ul style="list-style-type: none"> • Pressing • Gnawing • Crushing • Pinching • Cramping 	<ul style="list-style-type: none"> • Yes
Traction pressure		
<ul style="list-style-type: none"> • Tugging • Pulling • Wrenching 	<ul style="list-style-type: none"> • Tugging • Contraction (new) • Pulling • Wrenching • Clenching (new) 	<ul style="list-style-type: none"> • No
Thermal		
<ul style="list-style-type: none"> • Hot • Burning • Scalding (removed) • Searing 	<ul style="list-style-type: none"> • Hot • Searing • Burning 	<ul style="list-style-type: none"> • Yes
Brightness		
<ul style="list-style-type: none"> • Tingling • Itchy • Smarting (removed) • Stinging 	<ul style="list-style-type: none"> • Scratching (new) • Tingling • Itchy • Stinging 	<ul style="list-style-type: none"> • No
Dullness		
<ul style="list-style-type: none"> • Dull • Sore • Hurting • Aching • Heavy 	<ul style="list-style-type: none"> • Sore • Aching • Dull • Hurting • Heavy 	<ul style="list-style-type: none"> • Yes
Sensory miscellaneous		

Subclass and pain descriptors from the original MPQ	Suggested words for a new pain questionnaire	Ranking reordered? (yes or no)
<ul style="list-style-type: none"> • Tender • Taut (removed) • Rasping (removed) • Splitting 	<ul style="list-style-type: none"> • Tender • Splitting 	<ul style="list-style-type: none"> • No
Tension		
<ul style="list-style-type: none"> • Tiring • Exhausting 	<ul style="list-style-type: none"> • Tiring • Straining (new) • Exhausting 	<ul style="list-style-type: none"> • No
Autonomic		
<ul style="list-style-type: none"> • Sickening • Suffocating 	<ul style="list-style-type: none"> • Sickening • Suffocating 	<ul style="list-style-type: none"> • No
Fear		
<ul style="list-style-type: none"> • Fearful • Frightful (removed) 	<ul style="list-style-type: none"> • Fearful • Horrendous (new) • Horrifying (new) 	<ul style="list-style-type: none"> • No
Punishment		
<ul style="list-style-type: none"> • Terrifying • Punishing (removed) • Grueling • Cruel • Vicious • Killing 	<ul style="list-style-type: none"> • Grueling • Cruel • Vicious • Killing • Terrifying 	<ul style="list-style-type: none"> • Yes
Affective-evaluative-sensory miscellaneous		
<ul style="list-style-type: none"> • Wretched • Blinding 	<ul style="list-style-type: none"> • Wretched • Blinding 	<ul style="list-style-type: none"> • No
Evaluative		
<ul style="list-style-type: none"> • Annoying • Troublesome • Miserable • Intense • Unbearable 	<ul style="list-style-type: none"> • Mild (new) • Troublesome • Intense • Annoying • Irritating (new) • Unbearable • Horrible (new) • Miserable • Excruciating (new) • Distressing (new) 	<ul style="list-style-type: none"> • Yes
Supplementary a		
<ul style="list-style-type: none"> • Spreading • Radiating • Penetrating • Piercing 	<ul style="list-style-type: none"> • Spreading • Radiating • Penetrating • Piercing 	<ul style="list-style-type: none"> • No
Supplementary b		
<ul style="list-style-type: none"> • Tight • Numb • Drawing (removed) • Squeezing • Tearing 	<ul style="list-style-type: none"> • Bruising (new) • Tight • Numb • Squeezing • Tearing 	<ul style="list-style-type: none"> • No
Supplementary c		

Subclass and pain descriptors from the original MPQ	Suggested words for a new pain questionnaire	Ranking reordered? (yes or no)
<ul style="list-style-type: none"> Cool Cold Freezing 	<ul style="list-style-type: none"> Cool Cold Freezing 	<ul style="list-style-type: none"> No
Supplementary d		
<ul style="list-style-type: none"> Nagging Nauseating Agonizing Dreadful Torturing 	<ul style="list-style-type: none"> Nagging Nauseating Agonizing Dreadful Torturing 	<ul style="list-style-type: none"> No
Psychological (new)		
__ ^a	<ul style="list-style-type: none"> Worried (new) Angry (new) Fearful Sad (new) Depressed (new) Nervous (new) Anxious (new) Feel hopeless (new) Suicidal (new) 	<ul style="list-style-type: none"> N/A^b

^aNot available.

^bN/A: not applicable.

Removal of Words

Of the 78 pain descriptors, 8 (10%) pain descriptors from the original MPQ were removed because of low use (count<110). The words removed were *quivering* (from the *Temporal* subclass), *lancinating* (from the *Punctuate Pressure* subclass), *lacerating* (from the *Incisive Pressure* subclass), *scalding* (from the *Thermal* subclass), *smarting* (from the *Brightness* subclass), *taut* and *rasping* (from the *Sensory miscellaneous* subclass), and *frightful* (from the *Fear* subclass).

In addition, of the 78 pain descriptors from the original MPQ, 3 (4%) pain descriptors were removed because, on manual analysis of the social media posts, the contexts in which they were used were found to have deviated from pain description. The word *beating* (from the *Temporal* subclass) was mainly used as a synonym for *overcoming pain* or the physical action of a beating instead of being used as a pain descriptor. *Punishing* (from the *Punishment* subclass) and *drawing* (from the *Supplementary b* subclass) were also removed as these words were more often used as verbs rather than adjectives.

Addition of Words

A total of 13 new associated words were added to the pre-existing subclasses. The added words were *puncturing* (to the *Punctuate Pressure* subclass); *contraction* and *clenching* (to the *Traction Pressure* subclass); *scratching* (to the *Brightness* subclass); *straining* (to the *Tension* subclass); *horrendous* and *horrifying* (to the *Fear* subclass); *mild*, *irritating*, *horrible*, *excruciating*, and *distressing* (to the *Evaluative* subclass); and *bruising* (to the *Supplementary b* subclass).

Creation of a Psychological Subclass

An entirely new *Psychological* subclass was suggested to acknowledge the nonphysical aspects of pain. This subclass

comprised 9 newly identified descriptors relevant to the description of the emotional state of a person. These were selected for their high counts, which ranged from 1139 to 65,223. Intensity was also taken into consideration during this selection to ensure good graduality of pain description, and this ranged from 0.47 to 0.7467.

The identified pain descriptors consisted of a mixture of nouns and adjectives. For consistency, the 9 selected descriptors were modified to fit into the sentence “My pain makes me ____.” This resulted in the final selection of *worried*, *angry*, *fearful*, *sad*, *depressed*, *nervous*, *anxious*, *feel hopeless*, and *suicidal* for this new subclass.

Reordering of Pain Descriptors to Reflect Pain Intensity

The pain descriptors of 6 of the 20 subclasses were reordered based on their emotional intensity to reflect decreasing pain intensity. These new rankings are shown in [Table 1](#).

Discussion

Principal Findings

This study provides insights into modern language use in the context of pain description. We found infrequent use and even a change in context for the use of several descriptors from the original MPQ, reflecting the evolution of language and suggesting limitations in the current MPQ. We also identified several new pain descriptors that can be used to update the MPQ, including the emergence of a possible *Psychological* subclass of pain descriptors that can capture the emotional and mental aspects of pain. These changes may improve the application of the MPQ in physical settings and digital telehealth platforms. An updated MPQ with relevant and appropriate pain descriptors may improve the ability of the MPQ to accurately characterize the pain experience, leading to more tailored diagnoses and

treatment plans. In addition, better awareness of modern language use can also influence the design of conversational artificial intelligence systems, commonly known as *chatbots*, which have the potential to decrease the need for physical encounters in health care facilities [31]. The linguistic styles of chatbots are thought to affect relationship building between patients and the respective chatbots [32].

The top 10 highest counts for the 78 original MPQ keywords were substantially lower than those for the new words, ranging from 14,204 to 26,679 and from 21,441 to 96,909, respectively. This suggests that words previously selected for the MPQ may no longer be as commonly used today.

Interestingly, 6 out of the top 10 newly identified pain descriptors that received the top 10 highest counts were relevant to the emotional or mental description of pain, namely *anxiety*, *depression*, *fear*, *anger*, *depress*, and *sadness*, which led to the emergence of a *Psychological* subclass.

Our study combined the big data afforded by social media posts with artificial intelligence and additional emotion analytics algorithms that allow for a broad analysis of social media posts with high speed and accuracy. The combination of pain research with this technology, originally designed to help businesses understand what their customers want through linguistic and contextual cues, helps to address a pressing health care need. As much as 40% of the population contends with chronic pain, and the extensive cumulative impact of chronic pain in the United States alone was estimated to exceed US \$500 billion annually [33].

The updated pain definition by the International Association for the Study of Pain highlights the subjective and emotional nature of pain. The literature reports various links between pain and mood or psychiatric disorders. For instance, pain and major depressive disorder often occur concurrently, appearing to mutually exacerbate the severity of the individual conditions [34-37]. Chronic pain has been found to increase the risk of depressive disorders and comorbidity; inversely, depression is also a risk factor for the later development of chronic pain [38-42]. Although not fully understood, several studies have reported shared biological pathways and neurotransmitters between both conditions [43-45]. The involvement of the limbic structures of the brain in both pain processing and depressed mood can be triggered by stress and indicates the sharing of neurobiological factors between pain and depression [46-49]. In addition, the pain experienced may be altered by emotions such as anger expression [50], anxiety, or the feeling of powerlessness, with 1 author suggesting that reducing the perception of powerlessness may reduce pain intensity [51].

Unfortunately, these psychological, emotional, and behavioral interactions with pain experience are rarely brought up in patient interviews. The existing MPQ focuses largely on the physical description of pain. Although its existing words allow the interviewer to infer the emotional toll of pain on the patient, there is lack of a dedicated segment that acknowledges the psychological burden of pain. The introduction of a new

Psychological subclass is a significant change from the existing MPQ, prompting patients to definitively select the resultant emotion caused by their pain. The patient's choice from the selection of *worried*, *angry*, *fearful*, *sad*, *depressed*, *nervous*, *anxious*, *feel hopeless*, and *suicidal* may also indicate the patient's primary coping mechanism in the face of the challenges of the pain condition. The behavior of a patient who feels *anger* may be starkly different from another who feels *fearful* and could aid future interventions or referrals by health care providers.

Limitations and Future Direction

This study has some limitations. Only social media posts in English were included in this study, and most social media posts originated from the Western hemisphere. Therefore, the application of these findings is largely limited to the United States and may not be generalizable to other English-speaking countries. Future work using similar artificial intelligence mechanisms stratified by geographical regions to address regional linguistic differences, and perhaps different languages, may be explored.

Similarly, the sample population in this study was limited to users of social media platforms. The preferred social media platform differs according to age, and in general, the overall population tends to skew young [52,53]. It is possible that the pain descriptors found are specific to this population only; therefore, the results may not be generalizable to older adults for whom the older MPQ terminologies may still apply. There may also be a lack of representation of people of lower socioeconomic status or those whose pain conditions severely compromise daily living [54].

Although the artificial intelligence and emotion analytics algorithms selected for the study had previously demonstrated good accuracy, we did not evaluate the validity of the newly identified pain descriptors. The complexity of human thought and expression makes further validation of the suggested questionnaire on a sample of the target patients necessary. Future research could involve a qualitative focus group to examine the face validity of the word descriptors and to ensure that patients feel that their pain is adequately described by the options available or to suggest other terms. Validation and reliability of the new pain questionnaire should be conducted, and the proposed reordering of words within each subclass should be tested.

Conclusions

The original MPQ is inadequate for reporting the psychological aspects of pain. Several descriptors from the original MPQ were also noted to have infrequent use or changes in context. This study used artificial intelligence and emotion analytics algorithms to identify contemporary vocabulary for pain description. The described methodology could be repeated at regular intervals to ensure the relevance of the pain questionnaires. Further research is needed to examine the reliability and validity of the revised MPQ.

Acknowledgments

The authors would like to express their gratitude to the team at INTNT.AI (Singapore) for their invaluable support and technical guidance: Mr Manuel Ho, Mr Xuyuan Kee, Dr Murphy Choy, and Dr Yang Yinping. This study was supported by a grant from the Faculty of Dentistry, National University of Singapore.

Authors' Contributions

MYT, CEG, and HHT conceived and designed the study, collected the data, contributed to data or analysis tools (obtained through a partnership with INTNT.AI), and performed the analysis. MYT and CEG were involved with writing the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Pain descriptors for the first round of data gathering.

[\[DOCX File, 15 KB - formative_v5i11e31366_app1.docx\]](#)

Multimedia Appendix 2

List of common pain conditions.

[\[DOCX File, 14 KB - formative_v5i11e31366_app2.docx\]](#)

Multimedia Appendix 3

Counts and intensity of pain descriptors from the McGill Pain Questionnaire, web-based thesauruses, and identified through Word2Vec.

[\[DOCX File, 22 KB - formative_v5i11e31366_app3.docx\]](#)

References

1. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain* 2020 Sep 01;161(9):1976-1982 [FREE Full text] [doi: [10.1097/j.pain.0000000000001939](https://doi.org/10.1097/j.pain.0000000000001939)] [Medline: [32694387](https://pubmed.ncbi.nlm.nih.gov/32694387/)]
2. Eliav E, Gracely R. Measuring and assessing pain. In: Sharav Y, Benoliel R, editors. *Orofacial Pain and Headache*. Amsterdam: Elsevier; 2008:45-56.
3. Dansie E, Turk D. Assessment of patients with chronic pain. *Br J Anaesth* 2013 Jul;111(1):19-25 [FREE Full text] [doi: [10.1093/bja/aet124](https://doi.org/10.1093/bja/aet124)] [Medline: [23794641](https://pubmed.ncbi.nlm.nih.gov/23794641/)]
4. Breivik H, Borchgrevink P, Allen S, Rosseland L, Romundstad L, Hals EK, et al. Assessment of pain. *Br J Anaesth* 2008 Jul;101(1):17-24 [FREE Full text] [doi: [10.1093/bja/aen103](https://doi.org/10.1093/bja/aen103)] [Medline: [18487245](https://pubmed.ncbi.nlm.nih.gov/18487245/)]
5. Breivik EK, Björnsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. *Clin J Pain* 2000 Mar;16(1):22-28. [doi: [10.1097/00002508-200003000-00005](https://doi.org/10.1097/00002508-200003000-00005)] [Medline: [10741815](https://pubmed.ncbi.nlm.nih.gov/10741815/)]
6. Hicks C, von Baeyer CL, Spafford P, van Korlaar I, Goodenough B. The Faces Pain Scale - revised: toward a common metric in pediatric pain measurement. *Pain* 2001;93:173-183. [doi: [10.1016/S0304-3959\(01\)00314-1](https://doi.org/10.1016/S0304-3959(01)00314-1)]
7. McGuire DB. Comprehensive and multidimensional assessment and measurement of pain. *J Pain Sympt Manag* 1992 Jul;7(5):312-319. [doi: [10.1016/0885-3924\(92\)90064-o](https://doi.org/10.1016/0885-3924(92)90064-o)]
8. Clark C, Ferrer-Brechner T, Janal M, Carroll D, Yang J. The dimensions of pain: a multidimensional scaling comparison of cancer patients and healthy volunteers. *Pain* 1989;37(1):23-32. [doi: [10.1016/0304-3959\(89\)90149-8](https://doi.org/10.1016/0304-3959(89)90149-8)]
9. Clark WC, Carroll JD, Yang JC, Janal MN. Multidimensional scaling reveals two dimensions of thermal pain. *Pain* 1987;28(2):277-278. [doi: [10.1016/0304-3959\(87\)90143-6](https://doi.org/10.1016/0304-3959(87)90143-6)]
10. Dworkin R, Turk D, Revicki D, Harding G, Coyne K, Peirce-Sandner S, et al. Development and initial validation of an expanded and revised version of the Short-form McGill Pain Questionnaire (SF-MPQ-2). *Pain* 2009 Jul;144(1-2):35-42. [doi: [10.1016/j.pain.2009.02.007](https://doi.org/10.1016/j.pain.2009.02.007)] [Medline: [19356853](https://pubmed.ncbi.nlm.nih.gov/19356853/)]
11. Flaherty SA. Pain measurement tools for clinical practice and research. *AANA J* 1996 Apr;64(2):133-140. [Medline: [9095685](https://pubmed.ncbi.nlm.nih.gov/9095685/)]
12. Melzack R, Torgerson W. On the language of pain. *Anesthesiology* 1971 Jan;34(1):50-59 [FREE Full text] [doi: [10.1097/00000542-197101000-00017](https://doi.org/10.1097/00000542-197101000-00017)] [Medline: [4924784](https://pubmed.ncbi.nlm.nih.gov/4924784/)]
13. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1(3):277-299. [doi: [10.1016/0304-3959\(75\)90044-5](https://doi.org/10.1016/0304-3959(75)90044-5)]
14. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant

- Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken)* 2011 Nov 07;63(Suppl 11):240-252 [FREE Full text] [doi: [10.1002/acr.20543](https://doi.org/10.1002/acr.20543)] [Medline: [22588748](https://pubmed.ncbi.nlm.nih.gov/22588748/)]
15. Internet Live Stats. URL: <https://www.internetlivestats.com/> [accessed 2019-06-27]
 16. Hirschberg J, Manning CD. Advances in natural language processing. *Science* 2015 Jul 17;349(6245):261-266. [doi: [10.1126/science.aaa8685](https://doi.org/10.1126/science.aaa8685)] [Medline: [26185244](https://pubmed.ncbi.nlm.nih.gov/26185244/)]
 17. Subbiah IM, Hamilton E, Knoll M, Shanahan K, Meisel J. A big world made small: using social media to optimize patient care. *Am Soc Clin Oncol Educ Book* 2019 May;39:212-218. [doi: [10.1200/edbk.246643](https://doi.org/10.1200/edbk.246643)]
 18. McGloin AF, Eslami S. Digital and social media opportunities for dietary behaviour change. *Proc Nutr Soc* 2014 Oct 16;74(2):139-148. [doi: [10.1017/s0029665114001505](https://doi.org/10.1017/s0029665114001505)]
 19. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. *J Med Internet Res* 2014 Feb 11;16(2):e13 [FREE Full text] [doi: [10.2196/jmir.2912](https://doi.org/10.2196/jmir.2912)] [Medline: [24518354](https://pubmed.ncbi.nlm.nih.gov/24518354/)]
 20. Chew AM, Gunasekeran DV. Social media big data: the good, the bad, and the ugly (un)truths. *Front Big Data* 2021 Jun 1;4:623794 [FREE Full text] [doi: [10.3389/fdata.2021.623794](https://doi.org/10.3389/fdata.2021.623794)] [Medline: [34142082](https://pubmed.ncbi.nlm.nih.gov/34142082/)]
 21. Chan B, Lopez A, Sarkar U. The canary in the coal mine tweets: social media reveals public perceptions of non-medical use of opioids. *PLoS One* 2015 Aug 7;10(8):e0135072 [FREE Full text] [doi: [10.1371/journal.pone.0135072](https://doi.org/10.1371/journal.pone.0135072)] [Medline: [26252774](https://pubmed.ncbi.nlm.nih.gov/26252774/)]
 22. Reilly E, Sengupta R. Back pain, ankylosing spondylitis and social media usage; a descriptive analysis of current activity. *Rheumatol Int* 2020 Sep 19;40(9):1493-1499 [FREE Full text] [doi: [10.1007/s00296-020-04600-w](https://doi.org/10.1007/s00296-020-04600-w)] [Medline: [32430806](https://pubmed.ncbi.nlm.nih.gov/32430806/)]
 23. Ahlwardt K, Heavilin N, Gibbs J, Page J, Gerbert B, Tsoh JY. Tweeting about pain: comparing self-reported toothache experiences with those of backaches, earaches and headaches. *J Am Dent Assoc* 2014 Jul;145(7):737-743 [FREE Full text] [doi: [10.14219/jada.2014.30](https://doi.org/10.14219/jada.2014.30)] [Medline: [24982280](https://pubmed.ncbi.nlm.nih.gov/24982280/)]
 24. Wang D, Zou L, Pan W, Zhao D. Keyword graph: answering keyword search over large graphs. In: Zhou S, Zhang S, Karypis G, editors. *Advanced Data Mining and Applications*. Berlin, Heidelberg: Springer; 2012:635-649.
 25. Qiao M, Qin L, Cheng H, Yu JX, Tian W. Top-K nearest keyword search on large graphs. *Proc VLDB Endow* 2013 Aug 26;6(10):901-912. [doi: [10.14778/2536206.2536217](https://doi.org/10.14778/2536206.2536217)]
 26. Gupta R, Yang Y. CrystalNest at SemEval-2017 Task 4: using sarcasm detection for enhancing sentiment classification and quantification. In: *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. 2017 Presented at: 11th International Workshop on Semantic Evaluation (SemEval-2017); August, 2017; Vancouver, Canada p. 626-633. [doi: [10.18653/v1/s17-2103](https://doi.org/10.18653/v1/s17-2103)]
 27. Gupta RK, Yang Y. CrystalFeel at SemEval-2018 Task 1: understanding and detecting emotion intensity using affective lexicons. In: *Proceedings of The 12th International Workshop on Semantic Evaluation*. 2018 Presented at: The 12th International Workshop on Semantic Evaluation; June, 2018; New Orleans, Louisiana p. 256-263. [doi: [10.18653/v1/s18-1038](https://doi.org/10.18653/v1/s18-1038)]
 28. Haiman J. *Talk Is Cheap: Sarcasm, Alienation, and the Evolution of Language*. Oxford, United Kingdom: Oxford University Press; 1998:1-232.
 29. Crystalfeel - multidimensional emotion intensity analysis from natural language. A*STAR's Institute of High Performance Computing. 2021. URL: <https://socialanalyticsplus.net/crystalfeel> [accessed 2021-05-31]
 30. Goldberg Y, Levy O. word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv. 2014. URL: <https://arxiv.org/abs/1402.3722> [accessed 2021-10-30]
 31. Li LW, Chew AM, Gunasekeran DV. Digital health for patients with chronic pain during the COVID-19 pandemic. *Br J Anaesth* 2020 Nov;125(5):657-660 [FREE Full text] [doi: [10.1016/j.bja.2020.08.003](https://doi.org/10.1016/j.bja.2020.08.003)] [Medline: [32863018](https://pubmed.ncbi.nlm.nih.gov/32863018/)]
 32. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. *ACM Trans Comput-Hum Interact* 2005 Jun;12(2):293-327. [doi: [10.1145/1067860.1067867](https://doi.org/10.1145/1067860.1067867)]
 33. Institute of Medicine, Board on Health Sciences Policy, Committee on Advancing Pain Research, Care, and Education. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington, DC: National Academies Press (US); 2011.
 34. Li J. Pain and depression comorbidity: a preclinical perspective. *Behav Brain Res* 2015 Jan 01;276:92-98 [FREE Full text] [doi: [10.1016/j.bbr.2014.04.042](https://doi.org/10.1016/j.bbr.2014.04.042)] [Medline: [24797835](https://pubmed.ncbi.nlm.nih.gov/24797835/)]
 35. Ohayon MM. Specific characteristics of the pain/depression association in the general population. *J Clin Psychiatry* 2004;65(Suppl 12):5-9. [Medline: [15315471](https://pubmed.ncbi.nlm.nih.gov/15315471/)]
 36. Geerlings SW, Beekman AT, Deeg DJ, Twisk JW, Van Tilburg W. Duration and severity of depression predict mortality in older adults in the community. *Psychol Med* 2002 May 20;32(4):609-618. [doi: [10.1017/s0033291702005585](https://doi.org/10.1017/s0033291702005585)] [Medline: [12102375](https://pubmed.ncbi.nlm.nih.gov/12102375/)]
 37. IsHak WW, Wen RY, Naghdechi L, Vanle B, Dang J, Knosp M, et al. Pain and depression: a systematic review. *Harv Rev Psychiatry* 2018;26(6):352-363. [doi: [10.1097/hrp.000000000000198](https://doi.org/10.1097/hrp.000000000000198)]
 38. Hilderink P, Burger H, Deeg D, Beekman A, Voshaar RC. The temporal relation between pain and depression: results from the longitudinal aging study Amsterdam. *Psychosom Med* 2012;74(9):945-951. [doi: [10.1097/PSY.0b013e3182733fdd](https://doi.org/10.1097/PSY.0b013e3182733fdd)] [Medline: [23115345](https://pubmed.ncbi.nlm.nih.gov/23115345/)]

39. Yamamoto Y, Hayashino Y, Akiba T, Akizawa T, Asano Y, Saito A, et al. Depressive symptoms predict the subsequent risk of bodily pain in dialysis patients: Japan Dialysis Outcomes and Practice Patterns Study. *Pain Med* 2009 Jul 01;10(5):883-889. [doi: [10.1111/j.1526-4637.2009.00661.x](https://doi.org/10.1111/j.1526-4637.2009.00661.x)] [Medline: [19682272](https://pubmed.ncbi.nlm.nih.gov/19682272/)]
40. Fillingim RB, Ohrbach R, Greenspan JD, Knott C, Diatchenko L, Dubner R, et al. Psychological factors associated with development of TMD: the OPPERA prospective cohort study. *J Pain* 2013 Dec;14(12 Suppl):75-90 [FREE Full text] [doi: [10.1016/j.jpain.2013.06.009](https://doi.org/10.1016/j.jpain.2013.06.009)] [Medline: [24275225](https://pubmed.ncbi.nlm.nih.gov/24275225/)]
41. Knaster P, Karlsson H, Estlander A, Kalso E. Psychiatric disorders as assessed with SCID in chronic pain patients: the anxiety disorders precede the onset of pain. *Gen Hosp Psychiatry* 2012 Jan;34(1):46-52. [doi: [10.1016/j.genhosppsych.2011.09.004](https://doi.org/10.1016/j.genhosppsych.2011.09.004)] [Medline: [22001549](https://pubmed.ncbi.nlm.nih.gov/22001549/)]
42. McBeth J, Jones K. Epidemiology of chronic musculoskeletal pain. *Best Pract Res Clin Rheumatol* 2007 Jun;21(3):403-425. [doi: [10.1016/j.berh.2007.03.003](https://doi.org/10.1016/j.berh.2007.03.003)] [Medline: [17602991](https://pubmed.ncbi.nlm.nih.gov/17602991/)]
43. Gallagher R, Verma S. Managing pain and comorbid depression: a public health challenge. *Semin Clin Neuropsychiatry* 1999 Jul;4(3):203-220. [Medline: [10498788](https://pubmed.ncbi.nlm.nih.gov/10498788/)]
44. Blier P, Abbott F. Putative mechanisms of action of antidepressant drugs in affective and anxiety disorders and pain. *J Psychiatry Neurosci* 2001 Jan;26(1):37-43 [FREE Full text] [Medline: [11212592](https://pubmed.ncbi.nlm.nih.gov/11212592/)]
45. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. *Arch Intern Med* 2003 Nov 10;163(20):2433-2445. [doi: [10.1001/archinte.163.20.2433](https://doi.org/10.1001/archinte.163.20.2433)] [Medline: [14609780](https://pubmed.ncbi.nlm.nih.gov/14609780/)]
46. Giesecke T, Gracely RH, Williams DA, Geisser ME, Petzke FW, Clauw DJ. The relationship between depression, clinical pain, and experimental pain in a chronic pain cohort. *Arthritis Rheum* 2005 May;52(5):1577-1584 [FREE Full text] [doi: [10.1002/art.21008](https://doi.org/10.1002/art.21008)] [Medline: [15880832](https://pubmed.ncbi.nlm.nih.gov/15880832/)]
47. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo C, Kross E. An fMRI-based neurologic signature of physical pain. *N Engl J Med* 2013 Apr 11;368(15):1388-1397. [doi: [10.1056/nejmoa1204471](https://doi.org/10.1056/nejmoa1204471)]
48. Schmidt-Wilcke T, Clauw DJ. Fibromyalgia: from pathophysiology to therapy. *Nat Rev Rheumatol* 2011 Jul 19;7(9):518-527. [doi: [10.1038/nrrheum.2011.98](https://doi.org/10.1038/nrrheum.2011.98)] [Medline: [21769128](https://pubmed.ncbi.nlm.nih.gov/21769128/)]
49. Clauw DJ. Fibromyalgia: a clinical review. *J Am Med Assoc* 2014 Apr 16;311(15):1547-1555. [doi: [10.1001/jama.2014.3266](https://doi.org/10.1001/jama.2014.3266)] [Medline: [24737367](https://pubmed.ncbi.nlm.nih.gov/24737367/)]
50. Bruehl S, Chung OY, Burns JW. Anger expression and pain: an overview of findings and possible mechanisms. *J Behav Med* 2006 Dec 29;29(6):593-606. [doi: [10.1007/s10865-006-9060-9](https://doi.org/10.1007/s10865-006-9060-9)] [Medline: [16807797](https://pubmed.ncbi.nlm.nih.gov/16807797/)]
51. Walding MF. Pain, anxiety and powerlessness. *J Adv Nurs* 1991 Apr;16(4):388-397. [doi: [10.1111/j.1365-2648.1991.tb03427.x](https://doi.org/10.1111/j.1365-2648.1991.tb03427.x)] [Medline: [2061501](https://pubmed.ncbi.nlm.nih.gov/2061501/)]
52. Bolton RN, Parasuraman A, Hoefnagels A, Migchels N, Kabadayi S, Gruber T, et al. Understanding Generation Y and their use of social media: a review and research agenda. *J Serv Manag* 2013 Jun 14;24(3):245-267. [doi: [10.1108/09564231311326987](https://doi.org/10.1108/09564231311326987)]
53. Smith A, Anderson M. Social media use in 2018. Pew Research Center. 2018. URL: <https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/> [accessed 2021-10-30]
54. Caes L, Jones A, Jordan A. Engaging use of social media as a research tool to capture the daily life experiences of young people with chronic pain. *Evid Based Nurs* 2018 Oct 29;21(4):91-93. [doi: [10.1136/eb-2018-102980](https://doi.org/10.1136/eb-2018-102980)] [Medline: [30158175](https://pubmed.ncbi.nlm.nih.gov/30158175/)]

Abbreviations

MPQ: McGill Pain Questionnaire

Edited by G Eysenbach; submitted 18.06.21; peer-reviewed by D Gunasekeran, C Cheng, R Lee; comments to author 25.08.21; revised version received 06.09.21; accepted 18.09.21; published 25.11.21.

Please cite as:

Tan MY, Goh CE, Tan HH

Contemporary English Pain Descriptors as Detected on Social Media Using Artificial Intelligence and Emotion Analytics Algorithms: Cross-sectional Study

JMIR Form Res 2021;5(11):e31366

URL: <https://formative.jmir.org/2021/11/e31366>

doi: [10.2196/31366](https://doi.org/10.2196/31366)

PMID: [34842554](https://pubmed.ncbi.nlm.nih.gov/34842554/)

©Ming Yi Tan, Charlene Enhui Goh, Hee Hon Tan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.11.2021. 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 Patient-Generated Health Data From Twitter to Identify, Engage, and Recruit Cancer Survivors in Clinical Trials in Los Angeles County: Evaluation of a Feasibility Study

Katja Reuter^{1,2}, PhD; Praveen Angyan², MS; NamQuyen Le³, MPH; Thomas A Buchanan^{2,4}, MD

¹Department of Public Health and Preventive Medicine, The State University of New York Upstate Medical University, Syracuse, NY, United States

²Southern California Clinical and Translational Science Institute, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA, United States

³USC Annenberg School for Communication and Journalism, University of Southern California, Los Angeles, CA, United States

⁴Department of Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

Corresponding Author:

Katja Reuter, PhD

Department of Public Health and Preventive Medicine

The State University of New York Upstate Medical University

766 Irving Avenue

Syracuse, NY, 13210

United States

Phone: 1 3154641520

Fax: 1 3154641701

Email: reuterk@upstate.edu

Abstract

Background: Failure to find and attract clinical trial participants remains a persistent barrier to clinical research. Researchers increasingly complement recruitment methods with social media-based methods. We hypothesized that user-generated data from cancer survivors and their family members and friends on the social network Twitter could be used to identify, engage, and recruit cancer survivors for cancer trials.

Objective: This pilot study aims to examine the feasibility of using user-reported health data from cancer survivors and family members and friends on Twitter in Los Angeles (LA) County to enhance clinical trial recruitment. We focus on 6 cancer conditions (breast cancer, colon cancer, kidney cancer, lymphoma, lung cancer, and prostate cancer).

Methods: The social media intervention involved monitoring cancer-specific posts about the 6 cancer conditions by Twitter users in LA County to identify cancer survivors and their family members and friends and contacting eligible Twitter users with information about open cancer trials at the University of Southern California (USC) Norris Comprehensive Cancer Center. We reviewed both retrospective and prospective data published by Twitter users in LA County between July 28, 2017, and November 29, 2018. The study enrolled 124 open clinical trials at USC Norris. We used descriptive statistics to report the proportion of Twitter users who were identified, engaged, and enrolled.

Results: We analyzed 107,424 Twitter posts in English by 25,032 unique Twitter users in LA County for the 6 cancer conditions. We identified and contacted 1.73% (434/25,032) of eligible Twitter users (127/434, 29.3% cancer survivors; 305/434, 70.3% family members and friends; and 2/434, 0.5% Twitter users were excluded). Of them, 51.4% (223/434) were female and approximately one-third were male. About one-fifth were people of color, whereas most of them were White. Approximately one-fifth (85/434, 19.6%) engaged with the outreach messages (cancer survivors: 33/85, 38% and family members and friends: 52/85, 61%). Of those who engaged with the messages, one-fourth were male, the majority were female, and approximately one-fifth were people of color, whereas the majority were White. Approximately 12% (10/85) of the contacted users requested more information and 40% (4/10) set up a prescreening. Two eligible candidates were transferred to USC Norris for further screening, but neither was enrolled.

Conclusions: Our findings demonstrate the potential of identifying and engaging cancer survivors and their family members and friends on Twitter. Optimization of downstream recruitment efforts such as screening for *digital populations* on social media may be required. Future research could test the feasibility of the approach for other diseases, locations, languages, social media

platforms, and types of research involvement (eg, survey research). Computer science methods could help to scale up the analysis of larger data sets to support more rigorous testing of the intervention.

Trial Registration: ClinicalTrials.gov NCT03408561; <https://clinicaltrials.gov/ct2/show/NCT03408561>

(*JMIR Form Res* 2021;5(11):e29958) doi:[10.2196/29958](https://doi.org/10.2196/29958)

KEYWORDS

breast cancer; cancer; clinical research; clinical trial; colon cancer; inforeveillance; kidney cancer; lung cancer; lymphoma; patient engagement; prostate cancer; recruitment; Twitter; social media

Introduction

Background

Despite significant efforts to systematically describe barriers to identifying and enrolling clinical trial participants [1,2], insufficient recruitment of study participants remains a persistent roadblock to successful clinical research and medical progress [3-8]. A systematic review of randomized controlled trials found that 76.1% (131/172) of trials discontinued because of poor recruitment [9]. As many as 86% of clinical trials do not achieve accrual targets within their specified time [2,10]. Failure to find and attract eligible participants is a key factor contributing to clinical trial recruitment issues [9].

In recent years, researchers have increasingly complemented traditional recruitment methods (eg, flyers, public posters, advertisements in newspapers, radio, and television) with social media-based approaches [11,12]. Most of these studies have used either paid advertisements or organic, nonpaid messaging strategies to recruit study participants on social media. Social media also offers publicly accessible data from those who interact with and post on these platforms. This type of user-generated data can be used to rapidly capture and describe health-related attitudes and behaviors. Self-reported data from patients are referred to as patient-generated health data, that is, “health-related data created, gathered, or inferred by or from patients and for which the patient controls data collection and data sharing” [13].

Objectives

The goal of this pilot study is to examine the feasibility of using local user-reported data from cancer survivors and their family members and friends on the social network Twitter in Los Angeles (LA) County as a tool to enhance clinical trial recruitment at a comprehensive cancer center. According to the National Cancer Institute, “a person is considered a survivor from the time of diagnosis until the end of life” [14]. In this study, we included family members and friends because they play a critical role in caring for cancer survivors and in making cancer care decisions [15,16].

We used user-generated health data from the social network Twitter for 2 reasons: first, research has demonstrated the active use of Twitter among members of different patient communities who share their disease experiences, for example, cancer survivors [17-21], patients with diabetes [22], people with autism [23], and people with psoriasis [24]. Second, health surveillance researchers have demonstrated the usefulness of public Twitter data to understand public and patient perspectives

on a range of diseases and health topics, such as COVID-19, influenza, schizophrenia, smoking, HIV, and patient safety [25-32]. In some cases, social media user data have also demonstrated a correlation between disease prevalence and the frequency with which Twitter users discussed a disease [33].

This study focused on cancer survivors and their family members and friends who discussed any of the following 6 cancer conditions on Twitter in LA County (breast cancer, colon cancer, kidney cancer, lymphoma, non-small-cell lung cancer, and prostate cancer). We hypothesized that their user-generated data could be used to identify, engage, and potentially recruit cancer survivors for cancer trials.

Methods

Ethical Approval

This study relied on publicly available Twitter data. The authors adhered to Twitter’s terms of service and privacy policy [34,35]. Study-related data were collected using REDCap (Research Electronic Data Capture; Vanderbilt University) at the University of Southern California (USC), a secure, web-based application designed to support data capture for research studies. Any examples of Twitter account descriptions or tweets included in this report have been paraphrased to ensure user privacy. Study approval was obtained from the Clinical Investigations Committee at the USC Norris Comprehensive Cancer Center (USC Norris; Protocol 0S-17-7) and the Institutional Review Board at USC (Protocol HS-17-00811). This study was also registered at ClinicalTrials.gov (NCT03408561).

General Study Design and Study Setting

We set up the study as an interrupted time series with a before-and-after social media intervention.

The implementation site was USC Norris. Twitter data monitoring and data analysis were carried out at the School of Medicine at USC. The study protocol was published in the *Journal of Medical Internet Research Protocols* [36].

Intervention

Overview

The social media intervention to be tested in this study involved 2 steps: (1) monitoring cancer-specific posts about 6 cancer conditions (breast cancer, colon cancer, kidney cancer, lymphoma, lung cancer, and prostate cancer) posted by Twitter users in LA County with the goal of identifying cancer survivors and their family members and friends and (2) contacting eligible

users via public reply on Twitter to share information about related cancer trials that were open to accrual at USC Norris.

The intervention was used to identify potential trial participants (cancer survivors either directly or indirectly through their family members and friends) for all onboarded clinical trials. We refer to this approach as *centralized trial recruitment* because we clustered the trials into cancer disease groups and promoted only 6 disease trial groups on Twitter rather than each individual trial. This approach aligned with the USC Norris internal screening and triage process where physicians and clinical research coordinators are divided into disease-specific teams and thus will consider potential trial participants for all the relevant trials in the respective disease area. We onboarded one disease trial group every 2 weeks in a randomized order. The order in which the cancer trial disease groups were onboarded in this study was shuffled randomly using a Fisher-Yates shuffle [37]. Once a clinical trial disease group was onboarded, the trials in that group remained on for the period of this study.

Twitter Data Collection

To access public Twitter user data, we used Symplur Signals [38], a health care social media analytics platform that maintains a database of disease- and health-related Twitter posts and user data from the Twitter application programming interface. The messages that we extracted included at least one of the identified keywords related to the 6 cancer conditions of interest. The keywords and hashtags used in the search were determined using an iterative process based on an established conceptual framework for social data collection and quality assessment [39] and the Symplur Signals disease hashtag project [38]. The complete list of keywords used in the data search is listed in [Multimedia Appendix 1](#).

Twitter Data Verification and Analysis

We reviewed both retrospective and prospective data posted by Twitter users in LA County between July 28, 2017 and November 29, 2018. Because of the focus of the project, the analysis data set was limited to original tweets. Retweets (ie, Twitter posts shared by users who did not compose the original tweet) were removed from the data set. The goal was to identify cancer survivors and their family members and friends for each trial disease group. Twitter accounts and posts were reviewed and coded by 3 independent team members (KR, NL, and Alicia MacLennan). The coprincipal investigator reviewed any discrepancies to help resolve instances of disagreement between the coders.

We used a hybrid approach of qualitative research and machine learning (ML) methods to reliably determine the Twitter accounts of cancer survivors and their family members and friends. First, the classification of cancer survivors was based on self-reported information by Twitter users in either their profile description or tweet content. Data from users who did

not clearly state their cancer survivor status (or the cancer survivor status of a family member or friend) were excluded from the study. To do so, we manually reviewed each Twitter account (N=25,032), including the Twitter profile description (paraphrased example of a Twitter profile description of a cancer survivor: *Hard working man with big dreams, cancer survivor*) and the content from the tweets in our data set (paraphrased example of tweet from a family member of a cancer survivor: *Here's an update from my sister about her fight with stage IV colorectal cancer*). Only those data were included in the final data set in which users (cancer survivors or family members and friends) clearly mentioned their cancer experience or survivorship status (eg, on or off treatment). During the data review, we excluded non-English tweets from the analysis data set. In addition, 2 research team members independently reviewed the user accounts and coded the demographic characteristics of the Twitter users. We used 3 types of data for this analysis: (1) user profile description, (2) username, and (3) profile picture. Cases which were not clear were coded *unclear*. Because of the limited demographic information on Twitter, we limited this analysis to sex and race (people of color vs White).

Second, we used the ML program Brontometer (formerly BotOrNot) [40] established by Indiana University to verify our selection of human accounts. The program identifies automated Twitter accounts, the so-called bots [41], created by industry and interest groups that influence discussions and promote specific ideas or products [42,43]. Messages from these accounts pollute social and health research data sets. Botometer examines multiple variables such as the account's network (diffusion patterns), user (metadata), friends (account's contacts), temporal patterns (tweet rate), and sentiment (content of messages) and detects automated accounts with a 95% success rate [40].

Eligibility

Characteristics of Eligible Clinical Trials

Clinical trials were required to meet the eligibility criteria presented in [Textbox 1](#). Trial selection was independent of cancer stage. The 6 cancer trial disease categories were selected based on 2 factors: the results of a preliminary Twitter data analysis in California to determine the most frequently mentioned cancer topics in the region and the number of clinical trials at USC Norris that were open for accrual at the time. Between January 1, 2016 and January 30, 2017, we found 36,502 Twitter users in California who had sent a total of 159,396 Twitter messages in English including at least one of the selected 6 cancer disease terms (unpublished data from Symplur Signals). A preliminary analysis of clinical trials at USC Norris between January 1, 2017 and July 7, 2017 identified 84 clinical trials that were open for accrual at the time and were eligible for this study. We intended to onboard all eligible trials in the 6 cancer disease areas that were open for accrual at the time of the onset of this study.

Textbox 1. Inclusion and exclusion criteria for the clinical trials included in this study.

Inclusion criteria

1. Focus is on one of the 6 cancer disease types: breast cancer, colon cancer, kidney cancer, lymphoma, non–small-cell lung cancer, or prostate cancer.
2. Is Institutional Review Board–approved and open to accrual at the University of Southern California Norris Comprehensive Cancer Center.
3. Is a phase 1 trial in expansion, phase 2 or 3.
4. Is an interventional trial.
5. Has recruited in English
6. Has recruited for at least 9 months at enrollment.
7. Has set a monthly accrual target of ≥ 1 and annual accrual target of ≥ 12 .

Exclusion criteria

- Is a phase 1 trial in dose escalation.

Characteristics of Eligible Twitter Users

Participant recruitment for the onboarded clinical trials occurred on the social network Twitter. The study was limited to those Twitter users who met the eligibility criteria outlined in [Textbox 2](#). We applied both Boolean and Regex location code categories ([Multimedia Appendix 2](#)) to determine the user locations. Any

Twitter user located in LA County who mentioned at least one of the selected cancer disease topics ([Multimedia Appendix 1](#)) was contacted via Twitter using the public reply feature. We included all potential trial participants in this study who expressed interest in trial participation via Twitter or through the contact form on the trial webpage ([Figure 1](#)). They were invited to an initial phone prescreening.

Textbox 2. Inclusion and exclusion criteria for identifying eligible Twitter users.

Inclusion criteria

1. Is located in Los Angeles County based on the self-reported description provided on user’s Twitter profile.
2. Mentions in any of their Twitter messages at least one word or hashtag related to the 6 cancer disease types.
3. Message is an original Twitter message or reply to another user’s message
4. Message indicates that Twitter user has been diagnosed with cancer or that they know someone who has been diagnosed with cancer.

Exclusion criteria

1. Is younger than 18 years.
2. Notes that a relative or friend has died of the disease.
3. Retweets (ie, user shares message that other Twitter users sent).
4. Is a cancer survivor in remission with reduced signs and symptoms of the cancer.
5. Is a cancer survivor without any traces of cancer.

Figure 1. Example of a webpage that includes information about the clinical trials on lung cancer that are open to accrual at the University of Southern California (USC) Norris Comprehensive Cancer Center. Squares highlight the following page elements: a general video about reasons for participating in clinical trials, general information about clinical trials and USC Norris, and a privacy disclaimer including a URL link to the privacy policy of the USC. USC: University of Southern California.

General description of the purpose of these types of cancer trials

Study site information

Target audience

Contact form that triggers an email to this study team

List of all related open clinical trials at the USC Norris Comprehensive Cancer Center with a URL link to the study description on ClinicalTrials.gov for each trial

General video about reasons for participating in clinical trials

General information about clinical trials and the Center

Privacy disclaimer including a URL link to the privacy policy of the University of Southern California

Recruitment of Eligible Twitter Users

Any Twitter user who met the eligibility criteria (Textbox 2) was contacted via Twitter through the public message reply feature. No advertised (paid) messages were used in the outreach. The outreach messages (Textbox 3) were sent to each

user via the @USCTrials Twitter account. Each of the 3 messages served a different purpose in building trust and fostering investigator transparency [36,44]. The first message included a link to the related study webpage (Figure 1) that provided more information and allowed interested users to contact the trial coordination team.

Textbox 3. Outreach messages used to contact Twitter users.

Message type and the message

- Reply to the user's post that mentions the cancer disease including the link to related study webpage: "Dear (name): we noticed your mention of #kidneycancer and wanted to reach out. Did you know about these open #disease studies @KeckMedUSC? You can find more information here: (study page link) #ClinicalTrial"
- Message to introduce the research project: "We are reaching out to you as part of a research project trying to understand if Twitter can be used to better connect patients with clinical research opportunities."
- Privacy and security disclaimer: "The security of social media is not guaranteed. Contact us about the study. Do not post if concerned about privacy."

Participant Compensation

Participants (ie, contacted Twitter users) did not receive monetary compensation for their participation in the social media-based outreach but could receive compensation if they consented to participate in one of the clinical trials, depending on the trial.

Outcome Measures

The primary study outcomes included the proportion of Twitter users who (1) were identified on Twitter; (2) engaged with outreach messages through Twitter replies, mentions, likes, retweets, direct messages, or contact form use on the trial webpage; and (3) enrolled in a cancer trial. This paper reports the results of the tested intervention and related feasibility outcomes. It does not include the qualitative study results, that is, results of the qualitative interviews with study coordinators at the Cancer Center, to assess the feasibility and the level of acceptance that was also described in the protocol paper [36].

Analysis

Twitter Data Coding

All coding was performed independently by at least 2 research team members. The Cohen κ coefficient [45,46] was used to assess the interrater agreement of the 3 coders who analyzed

the Twitter accounts and posts. A coefficient greater than 0.8 was considered substantial for this study.

Statistical Analysis

As this was a pilot study, *P* values were of limited use to determine group differences. We used descriptive statistics to describe our study findings and calculated the proportions of targeted Twitter users who were identified, engaged with outreach messages, and enrolled in cancer trials.

Data Sharing

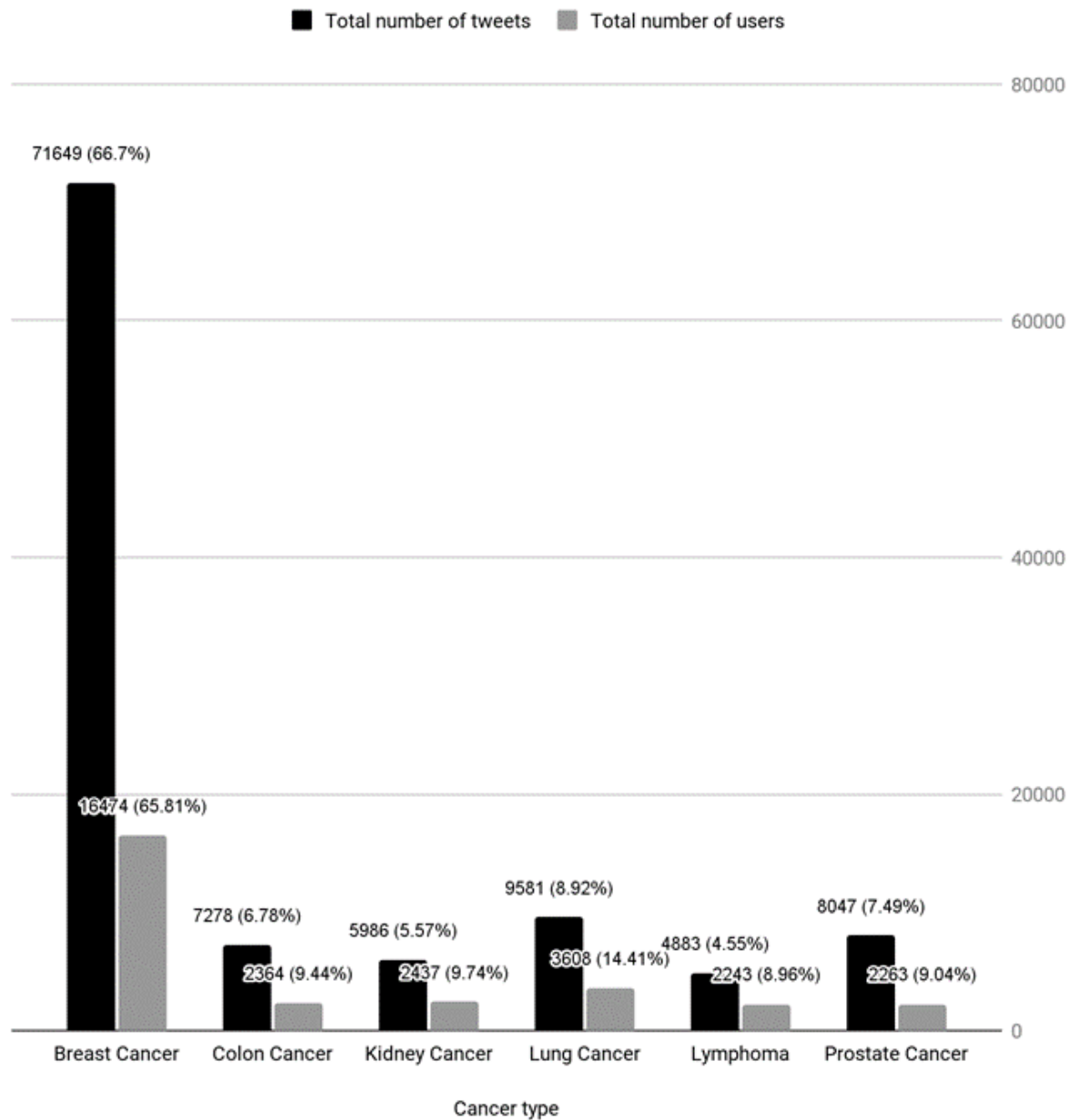
The deidentified and aggregated data that support the findings of this study are available in the data repository figshare [47-49].

Results

Volume of Cancer-Related Conversations on Twitter in LA County

Between July 28, 2017, and November 29, 2018, we retrieved and analyzed 107,424 tweets in English posted by 25,032 unique Twitter users in LA County (Figure 2). We found user-generated posts for all 6 cancer conditions. The highest number of tweets (71,649/107,424, 66.7%) and the highest number of users (16,474/25,032, 65.81%) were related to breast cancer (Figure 2). The topic lymphoma resulted in the lowest number of tweets (4883/107,424, 4.55%) and the lowest number of users (2243/25,032, 8.96%).

Figure 2. Number of Twitter posts and users in Los Angeles County by cancer type found between July 28, 2017, and November 29, 2018 in the data set of 107,424 posts.



Cancer Survivors and Their Family Members and Friends on Twitter in LA County

Of the 25,032 unique Twitter users in LA County who mentioned any of the cancer conditions, we identified 434 (1.73%) as eligible cancer survivors or their family members and friends. Approximately one-third (127/434, 29.3%) of them were cancer survivors, whereas the majority (305/434, 70.3%) mentioned a family member or friend who had cancer. Less than (2/434, 0.5%) discussed their own cancer experience and

the cancer of a family member or friend. In (4/434, 0.9%) of the posts, the person affected by cancer was unclear. Of the 434 users, 175 (40.3%) discussed breast cancer, 86 (19.8%) lymphoma, 64 (14.7%) lung cancer, 61 (14.1%) colon cancer, 28 (6.5%) prostate cancer, and 20 (4.6%) kidney cancer.

We further attempted to determine the sex and race of the eligible Twitter users (Table 1) and found that half of them were female, about one-third were male, and the sex of less than one-fifth was unclear. More than half of them were White, and about one-fifth were people of color.

Table 1. Demographics of the Twitter users who were identified as either cancer survivors or their family members and friends (n=434).

Characteristics	Participants, n (%)
Sex	
Male	140 (32.3)
Female	223 (51.4)
Unclear	71 (16.4)
Race	
People of color	99 (22.8)
White	254 (58.5)
Unclear	81 (18.7)

Responses to Cancer Trial Recruitment Messages on Twitter

Of the 434 targeted Twitter users we identified as eligible cancer survivors or their family members and friends, 85 (19.6%) engaged with the outreach messages (cancer survivors: 33/85, 38% and family members and friends: 52/85, 61%). We defined message engagement as any of the following user actions (on Twitter: reply, mention, *like*, retweet, and direct message and on the study webpage: contact form used to get in touch with the trial coordinator). In addition, of the 434 targeted Twitter users, 7 (1.6%) blocked the project account @USCTrials after the outreach, and 5 (1.2%) were deceased (based on replies by relatives or friends on Twitter). [Figure 3](#) presents the results of the main study procedures.

Positive user engagement regarding the intervention included feedback (paraphrased) such as:

I just finished my treatment, but this is such a good use of social media.

Do you know of programs or trials for colon cancer patients? I would like to learn if my family member is eligible for trials in the area.

However, it also became evident that some users included detailed medical information in their responses that posed user privacy risks, for example (paraphrased):

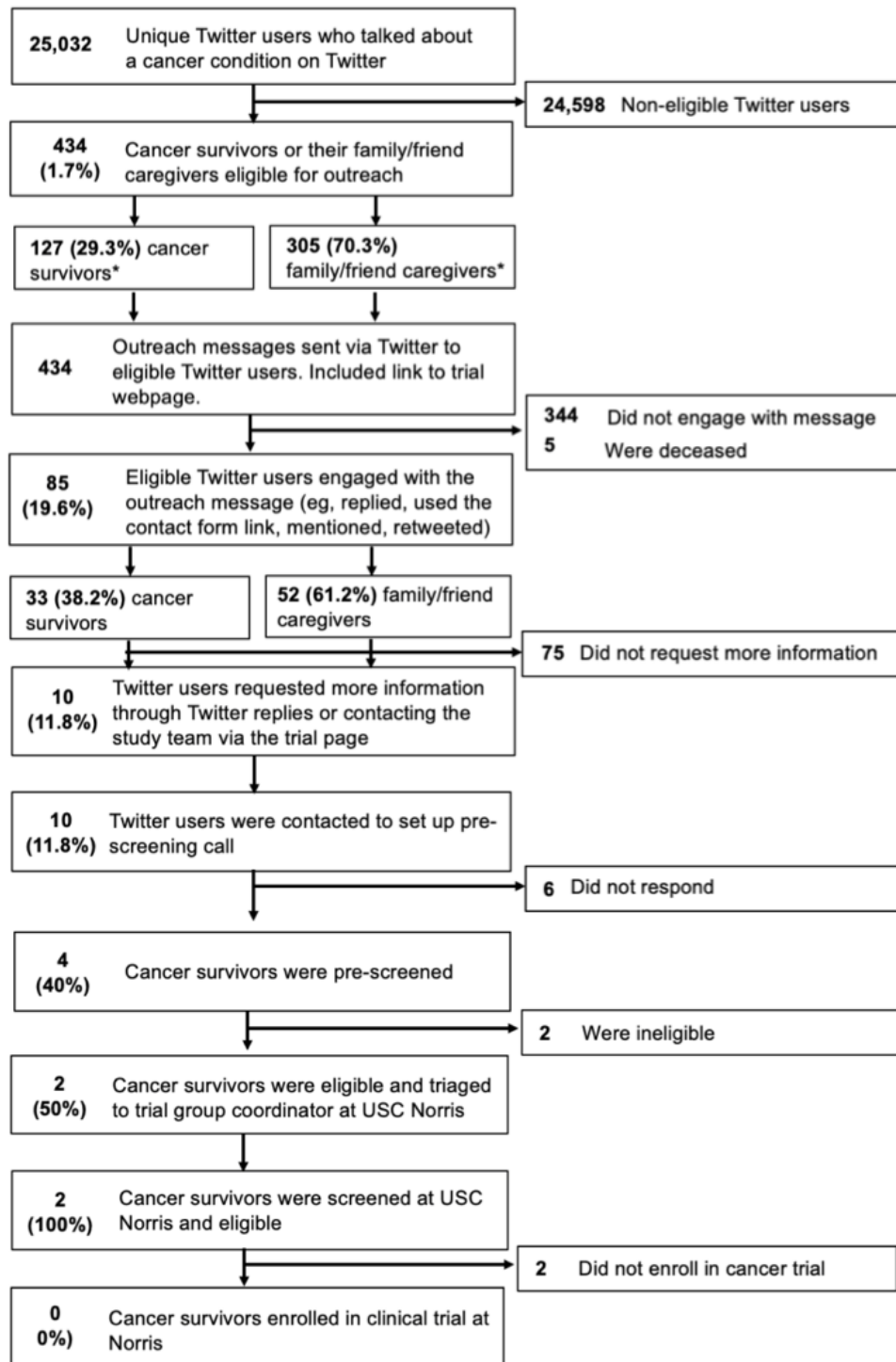
My wife is stage 4. They found lepto mets in her brain. After inserting an ommaya reservoir they started intrathecal chemotherapy, which didn't work so now they are considering partial brain radiation. She was also on letrozole and ibrance but they took her off due to her low blood counts. We don't favor brain radiation and would like another option. Do you have any related clinical trials?

Engagement with outreach messages varied according to cancer type. We found the highest level of engagement with the outreach messages for breast cancer (26/85, 30%), followed by lymphoma (22/85, 25%), lung cancer (18/85, 21%), colon cancer (12/85, 14%), prostate cancer (4/85, 4%), and kidney cancer (3/85, 3%).

We attempted to describe the characteristics of the 85 users who engaged with the recruitment messages. A quarter of them were male, whereas the majority were female. About one-fifth were people of color, whereas the majority were White ([Table 2](#)).

We further examined the engagement type with outreach messages by cancer survivors versus family members and friends ([Figure 4](#)). Of the 85 users who engaged with outreach messages, 33 (38%) were cancer survivors and 52 (61%) were family members and friends. Public replies (14/33, 42%), contact form use (11/33, 33%), and *likes* (14/33, 42%) were the primary forms of engagement among cancer survivors, whereas family members and friends primarily engaged via public reply (24/52, 46%) and *likes* (26/52, 50%).

Figure 3. Results for the main study procedures. USC Norris: University of Southern California Norris Comprehensive Cancer Center.

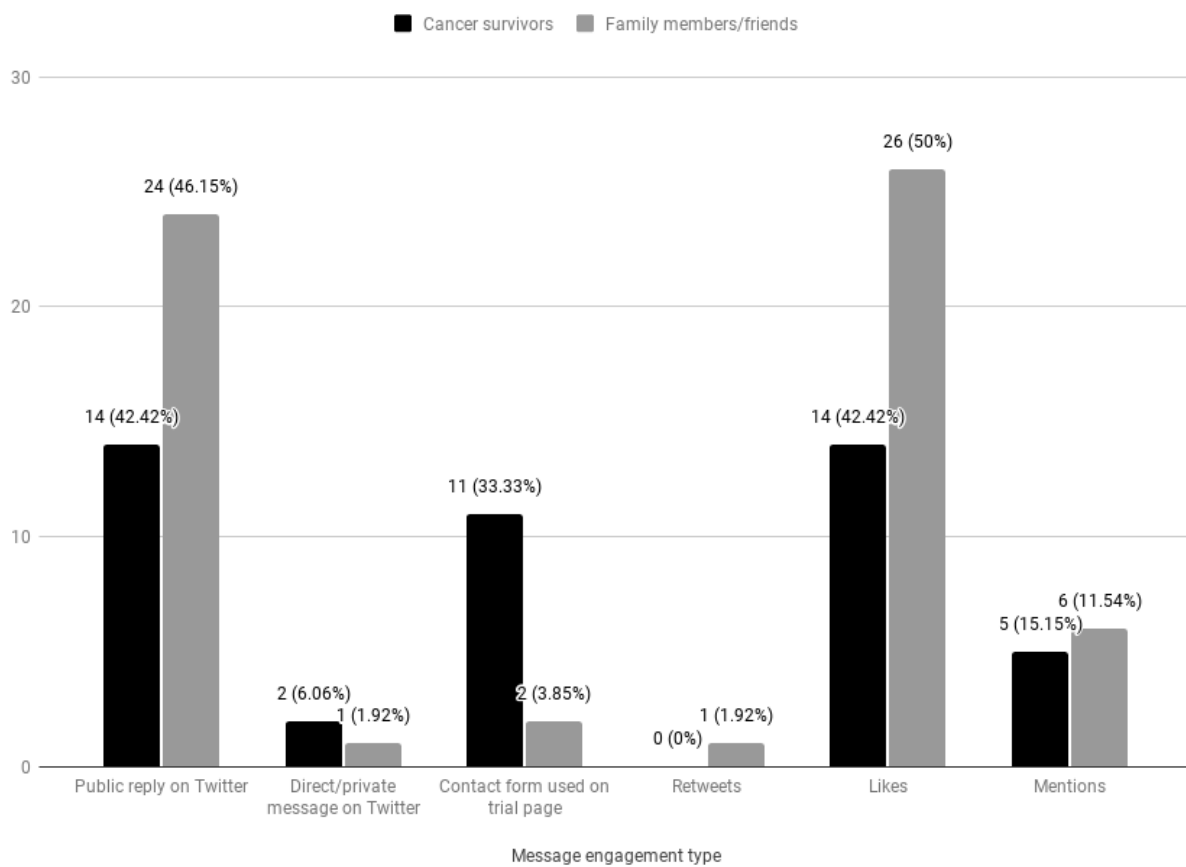


*Less than 0.5% (2/434) discussed their own cancer experience and the cancer of a family member or friend.

Table 2. Demographics of the Twitter users who engaged with the outreach messages (n=85).

Characteristics	Participants, n (%)
Sex	
Male	22 (25)
Female	52 (61)
Unclear	11 (12)
Race	
People of color	16 (18)
White	57 (67)
Unclear	12 (14)

Figure 4. Message engagement with outreach messages by cancer survivors versus family members and friends.



Prescreening, Screening, and Enrollment

Of the 85 Twitter users who engaged with the outreach messages, 10 (11%) requested more information through replies on Twitter, through emails, or through the use of the contact form on the trial webpage. Paraphrased example responses are as follows:

Thank you for sharing this information. How do I apply? I did not know about this work at USC.

I was diagnosed with colon cancer (stage 3) and had my operation in December. Now I'm on infusion

chemotherapy treatment, xelox and capecetabine. Please call me; I'm available by phone.

Of the 10 individuals who requested more information, 1 (10%) had breast cancer; 5 (50%), colon cancer; 1 (10%), kidney cancer; and 3 (30%), lymphoma. We followed up with these users either on Twitter or via email (in the case a user provided email information) to set up a prescreening phone call, but we only managed to set up a prescreening call with 40% (4/10) of them.

We triaged 2 eligible candidates—one with colon cancer and one with kidney cancer—and referred them to the USC Norris coordinator team for further screening. One participant qualified

for kidney cancer clinical trials, and the other qualified for colon cancer clinical trials. Finally, regardless of the abovementioned engagement with outreach messages, none of the recruitment outreach and engagement among Twitter users resulted in patients enrolling in a trial.

Discussion

Principal Findings

Our findings demonstrate that both cancer survivors and their family members and friends affected by different cancers can be found on Twitter based on their social media activity (ie, posts about cancer). In this study, approximately one-third of the identified Twitter users were cancer survivors who discussed their cancer experience, whereas others were family members and friends. Our data support the findings of previous studies that showed that both groups were present on social media for various reasons, such as seeking information resources and emotional and peer support on the web [17-21,50-52].

We found differences in the volume of posts and users across the 6 monitored cancer types. These variations could be explained by cancer prevalence and the size of the respective disease community on Twitter. For example, breast cancer is the most common type of cancer among women in the United States [53] and the third most common type in LA County [54]. It was also the most prevalent cancer topic in our data set (40.3% of the posts).

Our data further indicate that diverse, non-White cancer survivors and family members and friends can be found on social media, although more than half of the identified users were White. However, the social media user base has grown more representative of the broader population, and this trend is expected to continue. According to Pew Research [55], the percentage of American adults who use at least one social media site has increased across age groups: for example, among 18-29-year-olds, from 78% in 2009 to 90% in 2019; and among 50-64-year-olds, from 25% in 2009 to 69% in 2019. The same trend was reported for the racial diversity of social media users, but it may depend on the location in the United States. For example, based on data from 2019, on Twitter, 25% of the users are Hispanic, 24% Black, and 21% White.

Of the 434 targeted Twitter users we identified as eligible cancer survivors or their family members and friends, 85 (19.6%) engaged with the outreach messages. Some users included detailed medical information in their responses, which posed user privacy risks. Future studies are needed to develop strategies for managing such responses. We encouraged these Twitter users to delete their Twitter responses and continue the conversation with our team via the private direct message feature on Twitter, email, or phone.

Because of a lack of baseline data, it is challenging to determine whether we should have expected a higher message engagement rate. We directly contacted the cancer survivors or their family members and friends on Twitter who had mentioned the cancer keywords we were monitoring. We acknowledge that the level of acceptance of our direct targeting approach might vary among social media users. Gelinias et al [44] described this form of

recruitment as *active*, which “occurs when research staff approach and interact with specific individuals with the aim of enrolling them in research, usually on the basis of knowledge of characteristics that would make them suitable candidates for particular trials” [44]. However, before we implemented this pilot study, we examined the level of concern among Twitter users and nonusers about using Twitter surveillance data to recruit potential clinical trial participants [56]. We found that most Twitter users did not consider monitoring Twitter for clinical trial recruitment as inappropriate surveillance or a violation of privacy. That said, the expressed attitudes were highly contextual, depending on factors such as the type of disease or health topic (eg, HIV vs cancer vs smoking), the entity or person monitoring users on Twitter, and the monitored information. We found that the monitoring of Twitter user data related to HIV raised the highest level of concern compared with monitoring related to cancer, HPV, obesity, or smoking. This may be partly because HIV is still associated with stigma [57]. Bender et al [58] suggested that “within health information, there are gradients of sensitivity,” and certain health topics and disease types, such as cancer, may be considered less sensitive personal health information.

Furthermore, we did not tailor the content and language of the outreach messages to specific sex or racial and ethnic groups. Doing so may aid in engaging diverse social media users. One-fourth of those engaged with outreach messages were male, whereas more than half were female. About one-fifth were people of color, whereas more than half were White. Future research could assess the effectiveness of culturally tailored messaging strategies for this type of active recruitment.

Although one-fifth of the targeted Twitter users engaged with the outreach messages, our team was challenged to translate this type of user attention into prescreening phone calls, screening, and ultimate trial enrollment. We were cautiously optimistic that we would be able to enroll trial participants using this method. Our goal was to establish an effect size using the pilot data. However, we did not complete any enrollment. Nonetheless, we believe that the identification of cancer survivors and family members and friends creates opportunities for study recruitment and intervention research. Future research should examine the barriers to social media-based clinical trial recruitment.

The fact that less than half of the identified Twitter users responded to our efforts to set up a phone call indicates that successful social media engagement may not correlate with clinical trial enrollment. Without additional research into the barriers to social media-based clinical trial recruitment, it is difficult to identify adjustments required for the study design. Inputs from social media-based patient communities could shed light on the issue of the significant drop-off.

The barriers to follow through are still unknown and need to be studied. The lack of follow-through could be because of the variables that affect downstream clinical trial processes, such as screening and consenting. These aspects of the recruitment process are unrelated to social media monitoring and outreach. In addition, barriers can be at the patient level (eg, education about clinical trials and financial barriers), at the system level

(eg, inability to take time off work to schedule calls or appointments), or even at the community level (eg, lack of public transportation to the site and lack of resources) and can include factors related to the social determinants of health [7,59]. Approaches such as patient navigation have proven to be effective in addressing barriers at multiple levels in other settings [60] and could be adapted for social media-based recruitment methods. Digital forms of patient navigation could employ digital disease community members, influencers, and disease advocates on social media. We propose virtual focus groups and in-person interviews with cancer survivors on social media to identify barriers and facilitators of digital trial recruitment efforts and other types of interventions.

A detailed cost-effectiveness analysis was not included in this pilot study. However, information about the time and staff required may prove helpful in informing similar surveillance recruitment efforts. In this 1-year pilot, 4 team members were involved (10%-50% effort) to plan and implement the study (eg, developing the search strategy and weekly searches, coding extracted tweets and users to determine cancer survivors and family members and friends, sending outreach messages, tracking responses, following up and prescreening, and coordinating with the Cancer Center team). Future studies could conduct more robust cost-effectiveness and sustainability analyses.

Limitations and Challenges

We recognize that this is a pilot study and that the generalizability of the results is somewhat limited because cancer-related Twitter messages from locations outside of LA County and non-English messages (eg, Spanish) were not included. Moreover, social media interventions favor those with internet access and exclude those without, thus introducing bias in the participant population. The social media user base on Twitter is generally younger (38% are aged 18-29 years), college-educated, and located in urban or suburban areas [53] compared with the *average* study participant. Thus, recruiting via Twitter has the potential to select for this segment of the

population and could therefore introduce bias. Future research should determine the extension of the findings and conclusions to the population at large.

Much of the Twitter data we used were prospective. However, we also included retrospective social media data (ie, relevant Twitter messages sent by users in LA County 6 months before the onset of the study). Whether a user's message was posted 1 week or 6 months ago may have affected the user's engagement with the outreach messages. It is also possible that factors such as disease awareness months and trending news affect the attention to outreach messages.

Finally, the classification of cancer survivors was based on self-reported information from Twitter users. Although we excluded any data from users who did not clearly state their cancer survivor status (or the cancer survivor status of a family member or friend), there is a risk of misclassification, particularly if users provide false information. However, we are not aware of any research that suggests *patient imposters* on social media, that is, individuals falsely posing as patients.

Conclusions

Our findings demonstrate the potential of identifying and engaging cancer survivors and their family members and friends on Twitter, but we were unable to enroll clinical trial participants. However, we believe that the identification of cancer survivors and family members and friends creates opportunities for study recruitment and intervention research. Future research could examine barriers to this type of social media-based clinical trial recruitment, ways to tailor efforts downstream of the initial engagement, such as prescreening, screening, and consenting to the preferences of *digital populations* on social media, and the feasibility of the approach for other diseases, locations, languages, and social media platforms. Furthermore, the integration of computer science approaches such as deep learning, ML, and natural language processing to scale up the analysis of larger data sets would further support more rigorous testing of the intervention.

Acknowledgments

The authors acknowledge and thank Dr Anthony El-Khoueiry and Sarah Cole from the University of Southern California Norris Comprehensive Cancer Center for their inputs on the study design and help in aligning the study procedures with those at University of Southern California Norris. The authors also thank Alicia MacLennan for her help with data collection and cleaning. This work was supported by the Southern California Clinical and Translational Science Institute (grant UL1TR000130) from the National Center for Advancing Translational Sciences of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and National Center for Advancing Translational Sciences.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Cancer-specific keywords and hashtags were used to monitor Twitter user conversations in Los Angeles County. [[DOCX File, 25 KB - formative_v5i11e29958_app1.docx](#)]

Multimedia Appendix 2

Boolean and Regex location code categories were used in this study to determine a user's location in Los Angeles County. [[DOCX File , 24 KB - formative_v5i11e29958_app2.docx](#)]

References

1. Ross S, Grant A, Counsell C, Gillespie W, Russell I, Prescott R. Barriers to participation in randomised controlled trials: a systematic review. *J Clin Epidemiol* 1999 Dec;52(12):1143-1156. [doi: [10.1016/s0895-4356\(99\)00141-9](https://doi.org/10.1016/s0895-4356(99)00141-9)] [Medline: [10580777](https://pubmed.ncbi.nlm.nih.gov/10580777/)]
2. Treweek S, Lockhart P, Pitkethly M, Cook JA, Kjeldstrøm M, Johansen M, et al. Methods to improve recruitment to randomised controlled trials: Cochrane systematic review and meta-analysis. *BMJ Open* 2013;3(2):e002360 [FREE Full text] [doi: [10.1136/bmjopen-2012-002360](https://doi.org/10.1136/bmjopen-2012-002360)] [Medline: [23396504](https://pubmed.ncbi.nlm.nih.gov/23396504/)]
3. Treweek S, Pitkethly M, Cook J, Fraser C, Mitchell E, Sullivan F, et al. Strategies to improve recruitment to randomised trials. *Cochrane Database Syst Rev* 2018 Dec 22;2:MR000013. [doi: [10.1002/14651858.MR000013.pub6](https://doi.org/10.1002/14651858.MR000013.pub6)] [Medline: [29468635](https://pubmed.ncbi.nlm.nih.gov/29468635/)]
4. Bower P, Brueton V, Gamble C, Treweek S, Smith CT, Young B, et al. Interventions to improve recruitment and retention in clinical trials: a survey and workshop to assess current practice and future priorities. *Trials* 2014;15:399 [FREE Full text] [doi: [10.1186/1745-6215-15-399](https://doi.org/10.1186/1745-6215-15-399)] [Medline: [25322807](https://pubmed.ncbi.nlm.nih.gov/25322807/)]
5. Institute of Medicine. *Envisioning a Transformed Clinical Trials Enterprise in the United States: Establishing an Agenda for 2020: Workshop Summary*. Washington (DC): National Academies Press (US); 2012:1-248.
6. Institute of Medicine. *Public Engagement and Clinical Trials: New Models and Disruptive Technologies: Workshop Summary*. Washington (DC): National Academies Press (US); 2012.
7. Nipp RD, Hong K, Paskett ED. Overcoming barriers to clinical trial enrollment. *Am Soc Clin Oncol Educ Book* 2019 Jan;39:105-114 [FREE Full text] [doi: [10.1200/EDBK_243729](https://doi.org/10.1200/EDBK_243729)] [Medline: [31099636](https://pubmed.ncbi.nlm.nih.gov/31099636/)]
8. Huang GD, Bull J, Johnston McKee K, Mahon E, Harper B, Roberts JN, CTTI Recruitment Project Team. Clinical trials recruitment planning: a proposed framework from the Clinical Trials Transformation Initiative. *Contemp Clin Trials* 2018 Mar;66:74-79 [FREE Full text] [doi: [10.1016/j.cct.2018.01.003](https://doi.org/10.1016/j.cct.2018.01.003)] [Medline: [29330082](https://pubmed.ncbi.nlm.nih.gov/29330082/)]
9. Briel M, Olu KK, von Elm E, Kasenda B, Alturki R, Agarwal A, et al. A systematic review of discontinued trials suggested that most reasons for recruitment failure were preventable. *J Clin Epidemiol* 2016 Dec;80:8-15. [doi: [10.1016/j.jclinepi.2016.07.016](https://doi.org/10.1016/j.jclinepi.2016.07.016)] [Medline: [27498376](https://pubmed.ncbi.nlm.nih.gov/27498376/)]
10. Lamberti MJ, Mathias A, Myles JE, Howe D, Getz K. Evaluating the impact of patient recruitment and retention practices. *Drug Information J* 2012 Dec 30;46(5):573-580. [doi: [10.1177/0092861512453040](https://doi.org/10.1177/0092861512453040)]
11. Darmawan I, Bakker C, Brockman TA, Patten CA, Eder M. The role of social media in enhancing clinical trial recruitment: scoping review. *J Med Internet Res* 2020 Oct 26;22(10):e22810 [FREE Full text] [doi: [10.2196/22810](https://doi.org/10.2196/22810)] [Medline: [33104015](https://pubmed.ncbi.nlm.nih.gov/33104015/)]
12. 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 7;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/)]
13. Demiris G, Iribarren SJ, Sward K, Lee S, Yang R. Patient generated health data use in clinical practice: a systematic review. *Nurs Outlook* 2019;67(4):311-330. [doi: [10.1016/j.outlook.2019.04.005](https://doi.org/10.1016/j.outlook.2019.04.005)] [Medline: [31277895](https://pubmed.ncbi.nlm.nih.gov/31277895/)]
14. NCI dictionary of cancer terms. National Cancer Institute. 2011. URL: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/survivor> [accessed 2021-02-23]
15. Kim Y, Kashy DA, Kaw CK, Smith T, Spillers RL. Sampling in population-based cancer caregivers research. *Qual Life Res* 2009 Oct;18(8):981-989. [doi: [10.1007/s11136-009-9518-7](https://doi.org/10.1007/s11136-009-9518-7)] [Medline: [19655274](https://pubmed.ncbi.nlm.nih.gov/19655274/)]
16. Informal caregivers in cancer: roles, burden, and support (PDQ)—health professional version. National Cancer Institute. 2010. URL: <https://www.cancer.gov/about-cancer/coping/family-friends/family-caregivers-hp-pdq> [accessed 2021-03-26]
17. Modave F, Zhao Y, Krieger J, He Z, Guo Y, Huo J, et al. Understanding perceptions and attitudes in breast cancer discussions on Twitter. *Stud Health Technol Inform* 2019 Aug 21;264:1293-1297 [FREE Full text] [doi: [10.3233/SHTI190435](https://doi.org/10.3233/SHTI190435)] [Medline: [31438134](https://pubmed.ncbi.nlm.nih.gov/31438134/)]
18. Sedrak MS, Salgia MM, Bergerot CD, Ashing-Giwa K, Cotta BN, Adashek JJ, et al. Examining public communication about kidney cancer on Twitter. *JCO Clin Cancer Inform* 2019 Mar;3:1-6 [FREE Full text] [doi: [10.1200/CCL18.00088](https://doi.org/10.1200/CCL18.00088)] [Medline: [30860867](https://pubmed.ncbi.nlm.nih.gov/30860867/)]
19. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: example of cancer awareness via Twitter. *JMIR Public Health Surveill* 2016;2(1):e17 [FREE Full text] [doi: [10.2196/publichealth.5205](https://doi.org/10.2196/publichealth.5205)] [Medline: [27227152](https://pubmed.ncbi.nlm.nih.gov/27227152/)]
20. Sugawara Y, Narimatsu H, Hozawa A, Shao L, Otani K, Fukao A. Cancer patients on Twitter: a novel patient community on social media. *BMC Res Notes* 2012;5:699 [FREE Full text] [doi: [10.1186/1756-0500-5-699](https://doi.org/10.1186/1756-0500-5-699)] [Medline: [23270426](https://pubmed.ncbi.nlm.nih.gov/23270426/)]
21. Katz MS, Utengen A, Anderson PF, Thompson MA, Attai DJ, Johnston C, et al. Disease-specific hashtags for online communication about cancer care. *JAMA Oncol* 2016 Mar;2(3):392-394. [doi: [10.1001/jamaoncol.2015.3960](https://doi.org/10.1001/jamaoncol.2015.3960)] [Medline: [26539640](https://pubmed.ncbi.nlm.nih.gov/26539640/)]
22. Karahan I, Yürekli A, Özcömert OR, Oktaş B, Çifci A. Who tweets about diabetic foot on twitter and which tweets are more attractive? *Int J Low Extrem Wounds* 2020 Sep;19(3):251-254. [doi: [10.1177/1534734620912942](https://doi.org/10.1177/1534734620912942)] [Medline: [32204630](https://pubmed.ncbi.nlm.nih.gov/32204630/)]

23. Hswen Y, Gopaluni A, Brownstein JS, Hawkins JB. Using Twitter to detect psychological characteristics of self-identified persons with autism spectrum disorder: a feasibility study. *JMIR Mhealth Uhealth* 2019 Feb 12;7(2):e12264 [FREE Full text] [doi: [10.2196/12264](https://doi.org/10.2196/12264)] [Medline: [30747718](https://pubmed.ncbi.nlm.nih.gov/30747718/)]
24. Li W, Le N, Lee DJ, Reuter K. Analysis of psoriasis-related posts on Twitter: an abundance of patient-driven advocacy versus a scarcity of dermatologists. *J Am Acad Dermatol* 2020 Nov 07;S0190-9622(20):32914-32915. [doi: [10.1016/j.jaad.2020.11.004](https://doi.org/10.1016/j.jaad.2020.11.004)] [Medline: [33171165](https://pubmed.ncbi.nlm.nih.gov/33171165/)]
25. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. *PLoS One* 2020;15(4):e0231924 [FREE Full text] [doi: [10.1371/journal.pone.0231924](https://doi.org/10.1371/journal.pone.0231924)] [Medline: [32298385](https://pubmed.ncbi.nlm.nih.gov/32298385/)]
26. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: a surveillance study. *J Med Internet Res* 2020 Apr 09:e19016 [FREE Full text] [doi: [10.2196/19016](https://doi.org/10.2196/19016)] [Medline: [32287039](https://pubmed.ncbi.nlm.nih.gov/32287039/)]
27. Hswen Y, Naslund JA, Brownstein JS, Hawkins JB. Monitoring online discussions about suicide among Twitter users with schizophrenia: exploratory study. *JMIR Ment Health* 2018 Dec 13;5(4):e11483 [FREE Full text] [doi: [10.2196/11483](https://doi.org/10.2196/11483)] [Medline: [30545811](https://pubmed.ncbi.nlm.nih.gov/30545811/)]
28. Wagner M, Lampos V, Cox IJ, Pebody R. The added value of online user-generated content in traditional methods for influenza surveillance. *Sci Rep* 2018 Sep 18;8(1):13963 [FREE Full text] [doi: [10.1038/s41598-018-32029-6](https://doi.org/10.1038/s41598-018-32029-6)] [Medline: [30228285](https://pubmed.ncbi.nlm.nih.gov/30228285/)]
29. Nielsen RC, Luengo-Oroz M, Mello MB, Paz J, Pantin C, Erkkola T. Social media monitoring of discrimination and HIV testing in Brazil, 2014-2015. *AIDS Behav* 2017 Jul;21(Suppl 1):114-120 [FREE Full text] [doi: [10.1007/s10461-017-1753-2](https://doi.org/10.1007/s10461-017-1753-2)] [Medline: [28349220](https://pubmed.ncbi.nlm.nih.gov/28349220/)]
30. Malik A, Li Y, Karbasian H, Hamari J, Johri A. Live, love, juul: user and content analysis of Twitter posts about juul. *Am J Health Behav* 2019 Mar 01;43(2):326-336. [doi: [10.5993/AJHB.43.2.9](https://doi.org/10.5993/AJHB.43.2.9)] [Medline: [30808472](https://pubmed.ncbi.nlm.nih.gov/30808472/)]
31. Bychkov D, Young S. Social media as a tool to monitor adherence to HIV antiretroviral therapy. *J Clin Transl Res* 2018 Dec 17;3(Suppl 3):407-410 [FREE Full text] [Medline: [30873489](https://pubmed.ncbi.nlm.nih.gov/30873489/)]
32. Nakhasi A, Bell SG, Passarella RJ, Paul MJ, Dredze M, Pronovost PJ. The potential of Twitter as a data source for patient safety. *J Patient Saf* 2019 Dec;15(4):32-35. [doi: [10.1097/PTS.0000000000000253](https://doi.org/10.1097/PTS.0000000000000253)] [Medline: [26756726](https://pubmed.ncbi.nlm.nih.gov/26756726/)]
33. Tufts C, Polsky D, Volpp KG, Groeneveld PW, Ungar L, Merchant RM, et al. Characterizing tweet volume and content about common health conditions across pennsylvania: retrospective analysis. *JMIR Public Health Surveill* 2018 Dec 06;4(4):e10834 [FREE Full text] [doi: [10.2196/10834](https://doi.org/10.2196/10834)] [Medline: [30522989](https://pubmed.ncbi.nlm.nih.gov/30522989/)]
34. Twitter Terms of Service. URL: <https://twitter.com/en/tos> [accessed 2021-05-08]
35. Twitter Privacy Policy. URL: <https://twitter.com/en/privacy> [accessed 2021-05-08]
36. Reuter K, Angyan P, Le N, MacLennan A, Cole S, Bluthenthal RN, et al. Monitoring Twitter conversations for targeted recruitment in cancer trials in Los Angeles County: protocol for a mixed-methods pilot study. *JMIR Res Protoc* 2018 Sep 25;7(9):e177 [FREE Full text] [doi: [10.2196/resprot.9762](https://doi.org/10.2196/resprot.9762)] [Medline: [30274964](https://pubmed.ncbi.nlm.nih.gov/30274964/)]
37. Durstenfeld R. Algorithm 235: Random permutation. *Commun ACM* 1964 Jul;7(7):420. [doi: [10.1145/364520.364540](https://doi.org/10.1145/364520.364540)]
38. Disease hashtags. Symplur. URL: <https://www.symplur.com/healthcare-hashtags/diseases/> [accessed 2021-05-08]
39. Kim Y, Huang J, Emery S. Monitoring Twitter conversations for targeted recruitment in cancer trials in Los Angeles County: protocol for a mixed-methods pilot study. *J Med Internet Res* 2016;18(2):e41 [FREE Full text] [doi: [10.2196/jmir.4738](https://doi.org/10.2196/jmir.4738)] [Medline: [26920122](https://pubmed.ncbi.nlm.nih.gov/26920122/)]
40. Davis C, Varol O, Ferrara E, Flammini A, Menczer F. BotOrNot: a system to evaluate social bots. *arXiv cs.SI*. 2016. URL: <http://arxiv.org/abs/1602.00975> [accessed 2021-05-08]
41. Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. *Commun ACM* 2016 Jun 24;59(7):96-104 [FREE Full text] [doi: [10.1145/2818717](https://doi.org/10.1145/2818717)]
42. Mønsted B, Sapiezłyński P, Ferrara E, Lehmann S. Evidence of complex contagion of information in social media: an experiment using Twitter bots. *PLoS One* 2017;12(9):e0184148 [FREE Full text] [doi: [10.1371/journal.pone.0184148](https://doi.org/10.1371/journal.pone.0184148)] [Medline: [28937984](https://pubmed.ncbi.nlm.nih.gov/28937984/)]
43. Shao C, Ciampaglia GL, Varol O, Yang K, Flammini A, Menczer F. The spread of low-credibility content by social bots. *Nat Commun* 2018 Nov 20;9(1):4787 [FREE Full text] [doi: [10.1038/s41467-018-06930-7](https://doi.org/10.1038/s41467-018-06930-7)] [Medline: [30459415](https://pubmed.ncbi.nlm.nih.gov/30459415/)]
44. Gelinias L, Pierce R, Winkler S, Cohen IG, Lynch HF, Bierer BE. Using social media as a research recruitment tool: ethical issues and recommendations. *Am J Bioeth* 2017 Mar;17(3):3-14 [FREE Full text] [doi: [10.1080/15265161.2016.1276644](https://doi.org/10.1080/15265161.2016.1276644)] [Medline: [28207365](https://pubmed.ncbi.nlm.nih.gov/28207365/)]
45. Cohen J. A coefficient of agreement for nominal scales. *Edu Psychol Measur* 2016 Jul 02;20(1):37-46. [doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)]
46. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
47. Reuter K. Dataset: weekly tweet counts by cancer type per week. *Social and Community Informatics*. 2021. URL: https://figshare.com/articles/dataset/Tweet_counts_by_cancer_type_by_week/9461051 [accessed 2021-10-24]
48. Reuter K. Dataset: total tweet and user counts by cancer type. *Social and Community Informatics*. 2021. URL: https://figshare.com/articles/dataset/Total_tweet_and_user_counts_by_cancer_type/9461126 [accessed 2021-10-24]

49. Reuter K. Dataset: outreach message engagement. Social and Community Informatics. 2021. URL: https://figshare.com/articles/dataset/Outreach_message_engagement/14391752 [accessed 2021-10-24]
50. Daniels SR, Yang C, Toohey SJ, Willard VW. Perspectives on social media from adolescents and young adults with cancer. *J Pediatr Oncol Nurs* 2021 Mar 17;1043454221992319. [doi: [10.1177/1043454221992319](https://doi.org/10.1177/1043454221992319)] [Medline: [33729902](https://pubmed.ncbi.nlm.nih.gov/33729902/)]
51. Warner EL, Kirchoff AC, Wilson A, Cloyes KG, Sun Y, Waters AR, et al. Social support enactments on social media during the first 6 months of young adult cancer caregiving. *J Cancer Surviv* 2021 Mar 05:A. [doi: [10.1007/s11764-021-01004-y](https://doi.org/10.1007/s11764-021-01004-y)] [Medline: [33666872](https://pubmed.ncbi.nlm.nih.gov/33666872/)]
52. Booth A, Bell T, Halhol S, Pan S, Welch V, Merinopoulou E, et al. Using social media to uncover treatment experiences and decisions in patients with acute myeloid leukemia or myelodysplastic syndrome who are ineligible for intensive chemotherapy: patient-centric qualitative data analysis. *J Med Internet Res* 2019 Nov 22;21(11):e14285 [FREE Full text] [doi: [10.2196/14285](https://doi.org/10.2196/14285)] [Medline: [31755871](https://pubmed.ncbi.nlm.nih.gov/31755871/)]
53. Common cancer types. National Cancer Institute. 2015. URL: <https://www.cancer.gov/types/common-cancers> [accessed 2021-03-18]
54. Liu L, Wang Y, Sherman R, Cockburn M, Deapen D. Cancer in Los Angeles County: trends by race/ethnicity 1976-2012. Los Angeles Cancer Surveillance Program, University of Southern California. 2016. URL: <https://keck.usc.edu/cancer-surveillance-program/wp-content/uploads/sites/166/2016/08/Cancer-in-LA-2016.pdf> [accessed 2021-04-13]
55. Demographics of social media users and adoption in the United States. Pew Research Center: Internet, Science & Tech. 2019. URL: <https://www.pewinternet.org/fact-sheet/social-media/> [accessed 2019-06-28]
56. Reuter K, Zhu Y, Angyan P, Le N, Merchant AA, Zimmer M. Public concern about monitoring Twitter users and their conversations to recruit for clinical trials: survey study. *J Med Internet Res* 2019 Oct 30;21(10):e15455 [FREE Full text] [doi: [10.2196/15455](https://doi.org/10.2196/15455)] [Medline: [31670698](https://pubmed.ncbi.nlm.nih.gov/31670698/)]
57. Chambers LA, Rueda S, Baker DN, Wilson MG, Deutsch R, Raeifar E, et al. Stigma, HIV and health: a qualitative synthesis. *BMC Public Health* 2015 Sep 03;15:848 [FREE Full text] [doi: [10.1186/s12889-015-2197-0](https://doi.org/10.1186/s12889-015-2197-0)] [Medline: [26334626](https://pubmed.ncbi.nlm.nih.gov/26334626/)]
58. Bender JL, Cyr AB, Arbuckle L, Ferris LE. Ethics and privacy implications of using the internet and social media to recruit participants for health research: a privacy-by-design framework for online recruitment. *J Med Internet Res* 2017 Apr 06;19(4):e104 [FREE Full text] [doi: [10.2196/jmir.7029](https://doi.org/10.2196/jmir.7029)] [Medline: [28385682](https://pubmed.ncbi.nlm.nih.gov/28385682/)]
59. Marmot M. Social determinants of health: from observation to policy. *Med J Aust* 2000 Apr 17;172(8):379-382. [doi: [10.5694/j.1326-5377.2000.tb124011.x](https://doi.org/10.5694/j.1326-5377.2000.tb124011.x)] [Medline: [10840490](https://pubmed.ncbi.nlm.nih.gov/10840490/)]
60. Freeman HP. Patient navigation: a community based strategy to reduce cancer disparities. *J Urban Health* 2006 Mar;83(2):139-141 [FREE Full text] [doi: [10.1007/s11524-006-9030-0](https://doi.org/10.1007/s11524-006-9030-0)] [Medline: [16736361](https://pubmed.ncbi.nlm.nih.gov/16736361/)]

Abbreviations

LA: Los Angeles

ML: machine learning

REDCap: Research Electronic Data Capture

USC: University of Southern California

Edited by G Eysenbach; submitted 26.04.21; peer-reviewed by M Eder, J Benham, S Young; comments to author 01.06.21; revised version received 07.08.21; accepted 20.09.21; published 26.11.21.

Please cite as:

Reuter K, Angyan P, Le N, Buchanan TA

Using Patient-Generated Health Data From Twitter to Identify, Engage, and Recruit Cancer Survivors in Clinical Trials in Los Angeles County: Evaluation of a Feasibility Study

JMIR Form Res 2021;5(11):e29958

URL: <https://formative.jmir.org/2021/11/e29958>

doi: [10.2196/29958](https://doi.org/10.2196/29958)

PMID: [34842538](https://pubmed.ncbi.nlm.nih.gov/34842538/)

©Katja Reuter, Praveen Angyan, NamQuyen Le, Thomas A Buchanan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.11.2021. 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 for Using a Mobile App in Sickle Cell Disease Self-management: Descriptive Qualitative Study

Tilicia L Mayo-Gamble¹, MA, MPH, PhD; Delores Quasie-Woode², MPH, DrPH; Jennifer Cunningham-Erves³, MPH, MAEd, MS, PhD; Margo Rollins⁴, MD; David Schlundt⁵, PhD; Kemberlee Bonnet⁵, MA; Velma McBride Murry^{6,7}, PhD

¹Department of Health Policy and Community Health, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA, United States

²Center for Disease Control and Prevention Foundation, Atlanta, GA, United States

³Department of Internal Medicine, Meharry Medical College, Nashville, TN, United States

⁴Department of Pediatrics, Aflac Cancer and Blood Disorders Center, Emory University School of Medicine, Atlanta, GA, United States

⁵Department of Psychological Sciences, College of Arts and Sciences, Vanderbilt University, Nashville, TN, United States

⁶Department of Health Policy, Vanderbilt University School of Medicine, Nashville, TN, United States

⁷Department of Human and Organizational Development, Peabody College, Vanderbilt University, Nashville, TN, United States

Corresponding Author:

Tilicia L Mayo-Gamble, MA, MPH, PhD
Department of Health Policy and Community Health
Jiann-Ping Hsu College of Public Health
Georgia Southern University
501 Forest Dr
PO Box 8015
Statesboro, GA, 30458
United States
Phone: 1 912 478 1249
Email: tmayogamble@georgiasouthern.edu

Abstract

Background: Individuals with sickle cell disease (SCD) and their caregivers may benefit from technology-based resources to improve disease self-management.

Objective: This study explores the preferences regarding a mobile health (mHealth) app to facilitate self-management in adults with SCD and their caregivers living in urban and rural communities.

Methods: Five community listening sessions were conducted in 2 urban and rural communities among adults with SCD and their caregivers (N=43). Each session comprised 4 to 15 participants. Participants were asked questions on methods of finding information about SCD self-care, satisfaction with current methods for finding SCD management information, support for SCD management, important features for development of an mHealth app, and areas of benefit for using an mHealth app for SCD self-management. An inductive-deductive content analysis approach was implemented to identify the critical themes.

Results: Seven critical themes emerged, including the current methods for receiving self-management information, desired information, recommendations for communicating sickle cell self-management information, challenges of disease management, types of support received for disease management, barriers to and facilitators of using an mHealth app, and feature preferences for an mHealth app. In addition, we found that the participants were receptive to using mHealth apps in SCD self-management.

Conclusions: This study expands our knowledge on the use of mHealth technology to reduce information access barriers pertaining to SCD. The findings can be used to develop a patient-centered, user-friendly mHealth app to facilitate disease self-management, thus increasing access to resources for families of patients with SCD residing in rural communities.

(JMIR Form Res 2021;5(11):e28678) doi:[10.2196/28678](https://doi.org/10.2196/28678)

KEYWORDS

sickle cell disease; digital technology; rural; mHealth app; patient-centered technology; mobile health; health outcomes; hematology; mobile phone

Introduction

Background

Successful disease self-management improves health outcomes and decreases overall health care use [1,2]. For patients with chronic conditions, who are living in rural or medically underserved communities, self-management is essential because of their limited access to health care services and specialty care [3]. Technological advances present an opportunity for patients to access resources and health care providers to facilitate disease self-management. Past research has demonstrated that technology-based self-management tools are effective platforms for improving communication with health care providers, for patient education, for goal-setting and self-monitoring, for health care delivery, and for patient engagement [4-6]. Furthermore, health care organizations promote the use of technology-based interventions to minimize the disparities among vulnerable populations and extend the reach of care [7-12]. Technology-based interventions could potentially improve self-management in individuals with sickle cell disease (SCD), particularly for those residing in rural areas with limited access to specialty care (eg, hematology).

SCD is a blood disorder caused by the inheritance of 2 abnormal beta-globin alleles [13]. The most common form of SCD, both in the United States and worldwide, is homozygous hemoglobin SS, SCD/sickle cell anemia [14]. The disease affects approximately 100,000 individuals in the United States [15-17]. Individuals with SCD may experience acute and chronic complications because of ongoing vaso-occlusion crises secondary to hypoxia-induced sickling [18]. Acute pain is a hallmark of vaso-occlusion crises [19]. However, these individuals can develop stroke, priapism, and acute chest syndrome. Over time, acute vaso-occlusion episodes can result in chronic disease manifestations in end organs, such as poor renal, cardiac, pulmonary, and vascular dysfunction [20]. Therefore, disease management with a multidisciplinary team is crucial for managing SCD.

Although disease management is important, access to subspecialists is not readily available to all patients with SCD. Furthermore, education and interactive methods that prepare patients with SCD to implement self-care strategies to manage their disease are also limited. Therefore, having access to self-management tools could lead to a better understanding of disease management, increased adherence to self-care behaviors, and enhanced management of physical symptoms among individuals with SCD. Evidence of effective SCD self-management includes medication adherence, body temperature control, attention to appropriate food intake and hydration, and moderate exercising [21,22].

Objectives

Self-management programs incorporating mobile health (mHealth) apps are emerging and have demonstrated success

in self-care behaviors (eg, tracking pain, communicating with providers, and communicating with other individuals with SCD) in research studies targeting adolescents and young adults with SCD [23-25]. However, many of these studies did not include adults beyond the age of 22 years, who could also benefit from the mHealth resources. Furthermore, to our knowledge, there are no studies that have elicited input from patients on the usefulness of SCD mHealth apps for self-management, with a noticeable omission of patients and caregivers in rural communities [26]. These research gaps limit our insight into how mHealth apps could improve self-management among patients with SCD and caregivers, in general, and those residing in rural areas with limited access to subspecialty care, in particular. Tailoring health management technology to align with all needs of patients with SCD and preferences is necessary to increase access but may have implications for acceptance of these programs [27]. This study aims to contribute to the research and disease self-management practices of adults with SCD and their caregivers. Our sample included participants residing in urban and rural communities. The overall purpose of our study was to describe the preferences for mHealth app features designed to facilitate disease self-management.

Methods

Partnership

The researchers collaborated with the Sickle Cell Foundation of Tennessee (SCFT), a community-based organization serving patients with SCD and their caregivers. For this study, caregivers included the parents and spouses of individuals with SCD [28-30]. Representatives of the foundation were equal partners throughout each phase of the study, as described in the subsequent sections.

Recruitment and Study Population

The participants were recruited through the SCFT. Recruitment strategies included sending emails through the SCFT listserv and announcements made through SCFT support group meetings. Flyers were also distributed at hematology clinics in Memphis and Nashville, Tennessee. When participants contacted the organization or the researchers to express interest, they were asked about their city of residence. This self-reported information allowed us to ensure geographic diversity in the sample by including representation for participants from both urban and rural communities.

Five community listening sessions were conducted with adults having SCD and their caregivers in urban (3/5, 60%) and rural communities (2/5, 40%). A community listening session is a qualitative research method in which a group of people with diverse backgrounds and perspectives express their views on an important issue [31]. These sessions have demonstrated success because they are led by a trained community leader who is well-known within the target group and are less structured

than focus groups [31,32]. Our sessions ranged from 4 to 15 participants each (N=43).

Training

Before the community listening sessions, the research team conducted facilitator training using a manualized moderator's guide. Facilitators included the SCFT director and 2 community health workers who were well-known to the Tennessee SCD community through their advocacy. The manualized moderator's guide was developed jointly by the research team and the SCFT community partners. Members of the research team were PhD-level researchers with a qualitative research experience. In addition to receiving training in SCD management, the community listening session facilitators were trained in qualitative research as part of a larger patient engagement project [33]. Another aspect of the training included a 2-hour session on practical group engagement, which consisted of dyadic role playing, where the facilitators engaged in reflective techniques to encourage group discussion and establish group cohesion.

The questions posed during the community listening session focused on (1) methods to find information about SCD self-care, (2) satisfaction with current methods for finding SCD management information, (3) support for SCD management, (4) important features of an mHealth app, and (5) benefits of using an mHealth app for SCD self-management.

Procedures

Written informed consent was obtained from all the participants. The community listening sessions lasted approximately 2 hours and were audio-recorded. The facilitators explained the *ground rules*, which included the following statements: (1) everyone is encouraged to participate, (2) be respectful of other participants' responses, (3) there are no wrong or right responses, (4) one participant should respond at a time, and (5) feel free to ask questions anytime during the session. Before beginning the session, the participants completed a brief survey assessing their demographics and technology access. Each participant was reimbursed with US \$25 for partaking in the study. This study was approved by the institutional review board of Vanderbilt University.

Analysis

Coding and analysis were managed by the Vanderbilt University Qualitative Research Core, led by a PhD-level psychologist. Before the analysis, all sessions were commercially transcribed verbatim. An inductive-deductive content analysis approach was implemented to identify the critical themes [34]. The consolidated criteria for reporting qualitative research, an evidence-based qualitative methodology checklist, was used to guide the coding and analysis of the data [35]. Study questions and a preliminary review of transcripts were used as guides for developing a hierarchical coding system. A preliminary review of transcripts began after 2 listening sessions, allowing experienced coders to ensure reliability in using the coding system. Reliability is a process in which coders establish an agreement upon understanding the coding system [36]. Two coders encoded the same transcript and then compared and discussed any differences between their codes. The finalized code was used as data. After reaching consensus on the 2 transcripts, the subsequent transcripts were independently coded [36]. The coding of each transcript was compared, and any discrepancies were resolved to create a single-coded transcript. Each sentence was treated as a separate quote and assigned to 5 different codes. The transcripts were then combined and sorted using code. The analysis consisted of interpreting the coded quotes and identifying higher-order themes using an iterative inductive-deductive approach.

Results

Participant Characteristics

Table 1 presents the demographic characteristics of the participants. Most of the participants were women (30/43, 70%) with an average age of 45 years (SD 12.3), had an SCD genotype SS or SC (19/43, 44%), indicated their health status as fair to good (25/43, 58%), and had a household income of <US \$35,000 (25/43, 58%). Nearly half of the study participants (20/43, 47%) were unemployed. All the participants owned a smartphone to access mobile technology. Most participants had a smartphone with an iPhone (Apple Inc) operating system (25/43, 58%) and felt comfortable connecting to the internet on their smartphone (20/43, 47%). There were 67% (29/43) of participants who agreed that if an mHealth app offered information on SCD self-management, they would use it.

Table 1. Characteristics of patients and caregivers (N=43).

Characteristics	Values
Age (years), mean (SD)	44.8 (SD 12.3)
Sex, n (%)	
Female	30 (70)
Male	13 (30)
SCD^a genotype, n (%)	
Hb ^b SS ^c	10 (23)
Hb SC ^d	9 (21)
Hb Sβ0 ^e	1 (2)
Hb Sβ+ ^f	4 (9)
Unknown	4 (9)
Stakeholder group, n (%)	
Patient	36 (84)
Caregiver	7 (16)
Health insurance status, n (%)	
Uninsured	6 (14)
Private insurance	13 (30)
Medicaid or Medicare	19 (44)
Unknown	5 (12)
Perceived health status, n (%)	
Excellent	4 (9)
Very good	9 (21)
Good	10 (23)
Fair	15 (35)
Poor	1 (2)
Unknown	4 (9)
Education, n (%)	
Some HS ^g	3 (7)
High school diploma or GED ^h	4 (9)
Some college	17 (40)
College degree	13 (30)
Unknown	6 (14)
Marital status, n (%)	
Single	23 (54)
Married	10 (23)
Divorced	6 (14)
Unknown	4 (9)
Employment status, n (%)	
Unemployed	20 (47)
Employed (part time)	7 (16)
Employed (full time)	12 (28)
Unknown	4 (9)

Characteristics	Values
Household income level (US \$), n (%)	
≤10,000	8 (19)
10,001-25,000	12 (28)
25,001-35,000	5 (12)
35,001-45,000	4 (9)
>45,000	9 (21)
Unknown	5 (12)

^aSCD: sickle cell disease.

^bHb: hemoglobin.

^cHb SS: sickle cell disease or sickle cell anemia.

^dHb SC: sickle cell C disease.

^eHb Sβ0: sickle cell beta thalassemia zero.

^fHb Sβ+: sickle cell beta thalassemia plus.

^gHS: high school.

^hGED: general educational development.

In capturing the participants' perspectives, 7 themes emerged: (1) current methods for receiving self-management information; (2) desired information; (3) recommendations for communicating sickle cell self-management information; (4) challenges of disease management; (5) types of support received for disease management; (6) barriers to and facilitators of using an mHealth app; and (7) feature preferences for an mHealth app. For each major theme, 2 or more categories were identified and have been presented in [Textbox 1](#) along with illustrative quotes.

Textbox 1. Current methods for receiving self-management information. Illustrative quotes are examples from community listening participants used to develop the categories.

Illustrative quotes
<p>Health care providers:</p> <ul style="list-style-type: none"> • “The doctor is going to tell to take my medication.” [LS1] • “It’s through your doctors, is the main thing.” [LS1] • “Social media, the doctor.” [LS2] <p>Personal experience:</p> <ul style="list-style-type: none"> • “I’ve been going through it for so many years I kind of know what the doctors gonna tell me.” [LS1] • “The years of experience having it helps you to control it better.” [LS2] • “I’ve been going through this all my life so I know a pretty bit a much about it.” [LS3] • “A lot of learned things that we can, and can’t, or shouldn’t do through trial and error.” [LS4] <p>Internet:</p> <ul style="list-style-type: none"> • “If you’re not able to be with your doctor at that time where you can find something.” [LS1] • “I get a lot from social media.” [LS2] • “On the internet.” [LS3] • “We used the websites in early days.” [LS4] <p>Other individuals with sickle cell disease:</p> <ul style="list-style-type: none"> • “Getting our men to talk together.” [LS2] • “Talking to each other, it’s like we’re the same people sometimes because we all go through the same thing, and we bounce information off of one another.” [LS4] • “My grandmother already had it, so we already understood how we was going to get through this.” [LS4]

Theme 1: Current Methods for Receiving Self-management Information

This theme explained how the participants learned about the various methods of disease management. We identified 4 categories of these methods: health care providers, personal experience, internet, and other individuals with SCD. Health care providers were the primary source of self-management information. The participants admitted to becoming more knowledgeable about SCD and managing its symptoms through their health care providers, “The doctor is going to tell to take my medication” (LS1). However, because every patient is different, participants also indicated that they relied on their personal lived experiences to guide self-management. For example, one participant stated, “The years of experience having it helps you to control it better” [LS2]. Participants also mentioned learning about sickle cell (disease) from other individuals with SCD. This appears to be an important method for information that resonates among participants; “Talking to each other, it’s like we’re the same people sometimes because we all go through the same thing” [LS4]. In addition,

participants stated that the internet was a source of learning about self-management techniques. Specific internet sites included finding information on social media. One participant stated, “I got a lot from social media” [LS2].

Theme 2: Desired Information

The *desired information* theme reflected the areas in which the participants wanted more knowledge (Textbox 2). The 2 categories of information were SCD and SCD treatment. Participants expressed a desire to learn more about the differences in symptoms among males and females, along with steps to manage their care: “I would recommend something along the lines of distinguishing protocols and forces of action for male and female” [LS1]. Some other participants had questions about disease management. For example, one participant indicated, “Do’s and don’ts. Yeah, just some basic lists, create a list of those types of things to share and communicate” [LS4]. Overall, despite receiving care for SCD, knowledge gaps still existed between the disease and the options to facilitate its self-management.

Textbox 2. Desired information. Illustrative quotes are examples from community listening participants recruited to develop the categories.

Illustrative quotes

Sickle cell disease (eg, types and symptoms):

- “Any other stress factors or different things like that, that I think are different between men and women.” [LS1]
- “I can concur there is a big dip there (surrounding SCD) for the college students struggling to figure out.” [LS2]
- “Direct sickle cell. Like you said it’s different ones. The ones pertaining to yourself.” [LS3]
- “I just wonder why it’s not out there a bit more about the trait that has symptoms.” [LS3]

Treatment:

- “The biggest gap that I have seen in coming from the standpoint of the patient, is a transitional period.” [LS2]
- “In Memphis about treatment about possible cure for sickle cell disease.” [LS3]
- “New studies, new medications.” [LS4]
- “I think that it’s kind of going back to that point where I think about some of the cancer examples.” [LS5]
- “I’ve heard about where they’re super engaged. They know about this study being done over there.” [LS5]

Theme 3: Recommendations for Communicating Sickle Cell Self-management Information

This theme describes the participants’ reflections on strategies to increase their knowledge of self-management (Textbox 3). The categories within this theme included suggestions on how the information could be improved, how the information was used in daily disease management, and the preferred methods for receiving the information. The participants agreed that communicating information within a community was important. For example, one participant stated, “I think when you go back out to these communities there should be probably a stronger dialogue between where are we in terms of what can we do and what they want to see done” [LS5]. Regarding the use of

information for daily disease management, a participant stated, “We have to kind of construct the feedback you get from the general population in a way that we can act on” [LS5]. For the preferred methods for receiving information, the participants discussed tailoring of information. For example, “I think that part of the trick is finding out how those are incorporated into how a person goes through their life” [LS5]. Both categories seemed to indicate a need for more specific information that may be incorporated into daily disease self-management. Although many discussed using the internet to receive information, one participant also acknowledged the importance of finding reputable sites, “They taught you to actually go to reputable sites and look for stuff because we know that anybody can put stuff on the internet” [LS2].

Textbox 3. Recommendations for communicating sickle cell self-management information. Illustrative quotes are examples from community listening participants recruited to develop the categories.

Illustrative quotes

How information could be improved:

- “The biggest gap that I’ve seen in coming from the standpoint of the patient, is a transitional period. It’s either pediatrics or they’re adults.” [LS2]
- “You’re right. Stop telling these kids they not going to live long.” [LS4]
- “I think when you go back out to these communities there should be probably a stronger dialogue between where are we in terms of what can we do and what they want to see done.” [LS5]

How information is used for daily disease management:

- “And they can email me that list or it could show up somewhere for me to access it.” [LS1]
- “We have to kind of construct the feedback you get from the general population in a way that we can act on.” [LS5]

Preferred methods for receiving information:

- “And it would be in my phone so the next time I go, and be like, you know, you see this, you know.” [LS1]
- “Put it on the phones so I can get a reminder. Put on the calendar, whatever.” [LS4]
- “I think that part of the trick is finding out how those are incorporated into how a person goes through their life.” [LS5]

Theme 4: Challenges of Disease Management

This theme captures the difficulties experienced by the participants in maintaining a healthy lifestyle with SCD. Categories included limited self-care, limited provider knowledge on SCD, provider disrespect or dismissiveness, inadequate care, and the impact of disease on personal relationships. Limited provider knowledge on SCD is illustrated by the following quote, “They went to school and learned from books, but they can’t know everything, you know” [LS3]. Many participants expressed that they were displeased by the responses from the providers when they asked questions about managing their disease. Participants also discussed the experience of

feeling dismissed or disrespected by a provider, “You gonna get in and they ain’t going to listen to nothing you tell them” [LS1]. Participants also felt the care they received from the providers was inadequate, “It don’t matter the disease, sickle cell, we still treated like some of the bottom” [LS3]. As indicated in [Textbox 4](#), this treatment was often experienced in emergency room settings. In addition, the participants described the toll that the disease management took on their personal relationships. One participant stated, “I can take a big amount of pain, so I try to hide my pain because it’s people that depend on me and I don’t be wanting to let them down” [LS3]. Participants also shared that they “hide their pain because they do not want others to feel sorry for them.”

Textbox 4. Challenges of disease management. Illustrative quotes are examples from community listening participants recruited to develop the categories.

<p>Illustrative quotes</p> <p>Limited self-care:</p> <ul style="list-style-type: none"> • “It’s like I don’t care how intense the pain is, I’ll check and see how I feel tomorrow. I’ll check tomorrow. If I put it off hopefully I’ll start feeling better.” [LS2] • “Partying, drinking knowing that you can dehydrates us and put us in a crisis. Just doing everything that we shouldn’t be doing.” [LS4] • “If I don’t take my medicine on a daily basis it because I don’t feel good and I go through that blood issue a lot and I don’t feel good.” [LS4] • “I’m just stubborn you know? My wife has to make me go to the hospital.” [LS1] <p>Limited provider knowledge on sickle cell disease:</p> <ul style="list-style-type: none"> • “I could remember at [...] State, and being in the health center and they didn’t have a clue as to, you know, what to do.” [LS1] • “Yeah, more than my doctors. They think they know everything.” [LS3] • “What may work for me doesn’t work for you. What may work for him, doesn’t work for her. The doctors need to realize that.” [LS4] • “When I was sick as a child. They had no idea what sickle cell was in the town that I grew up in.” [LS4] • “Right now sickle cell is one page in their medical book.” [LS4] • “I feel like I’m treated better when somebody is with me. when you go to the ER with them, the time that it takes for them to get you back, and get your information, and to see if it’s okay to get medication, or IV.” [LS4] • “Sickle cell, we still treated like some of the bottom. No matter the disease period, it has to start somewhere.” [LS4] • “There’s going to be ignorance towards the disease, you’ll just so happen to get most of it and it’s frustrating.” [LS4] • “Lack of education of providers around SCD.” [LS5] <p>Provider disrespect or dismissiveness:</p> <ul style="list-style-type: none"> • “Like you said we tell them it’s a 10, and they’re like ‘Are you sure?’ You know, you don’t, they don’t say ‘Are you sure it’s a 10?’” [LS1] • “And I’ve had nurses try to, you know, you know, stick me. And I tell them, you know, ‘It’s gonna be hard.’ Oh no, and then they missed it.” [LS1] • “I had an episode where I went to the hospital and he came in, listened to me, ask me what was going on, walked out the room, the nurse came back in and said, ‘He’s going to send you home.’ Didn’t do nothing.” [LS4] • “I can tell them, ‘This is what I need you to do, Do A, B, C, D.’ And guess what he does, D, F, G. He ain’t heard what I said. He ain’t heard nothing I just told him.” [LS4] • “Because they get abused in the system. ‘Why are you here? Is this any different than before?’” [LS5] <p>Inadequate care:</p> <ul style="list-style-type: none"> • “But, if I get to the hospital, and they’re not giving me any more than I can already take at home, I’m getting frustrated.” [LS1] • “You find it more on children but adult, you’re gonna be in the emergency room two or three hours before you get any help anyway and you might’ve could’ve stayed at home and got yourself together a little bit more, you know?” [LS2] • “You gonna get in and they ain’t going to listen to nothing you tell them, so it’s like you was fighting a no win situation. [LS3] • “I had a doctor walk in the room. I’m in full blown crisis. I’m sweating, everything going crazy. He walks in the room, look at me, “What’s wrong?” Did tell him what’s wrong. He walked right back out the room and never came back.” [LS4] • “Even when we have therapy in general for pain management, we only have four slots a day for 260 patients, and that’s not guaranteed. That’s also a burden to kind of try to fix what we actually can.” [LS5] <p>Impact of disease on personal relationships:</p> <ul style="list-style-type: none"> • “Cause you go from spouse to mother.” [LS1] • “With me, I can take a big amount of pain so I try to hide my pain because it’s people that depend on me and I don’t be wanting to let them down.” [LS3]

Theme 5: Types of Support Received for Disease Management

This theme focuses on interpersonal relationships that can support maintaining a healthy lifestyle (Textbox 5). The

categories of support that were discussed included emotional and informational support from family, health care accompaniment, and practical assistance with physical care. Participants explained the benefit of familial social support including, “I typically will call people that I know in my

network, that understand the disease. Most of my friends, most of my family, people that I know, I've pretty much got them educated on it" [LS4]. This also suggests that SCD self-management may require patients to act as educators for

members of their support network. The participants also discussed examples of families helping with physical care. For example, one participant stated, "People try to help with the massages and all" [LS3].

Textbox 5. Types of support received for disease management. Illustrative quotes are examples from community listening participants recruited to develop the categories.

Illustrative quotes

Familial:

- "Helping me talk to the doctor. 'Well what happen if she has to take this, this, and this?'" [LS1]
- "At this point it's me and my husband understands. When I was younger it was my mom, my grandmother, my dad." [LS4]
- "Yeah. Poor performing physicians get a lower salary reimbursement from the hospital, so they started paying attention to these families, but we have to empower the families to know how to complain." [LS5]

Health care accompaniment:

- "Definitely nurse or advocate because my mom used to be a nurse at Vanderbilt, so. When it was her, she would go into nurse mode." [LS1]
- "It's cold, I'm in pain and I'm mad and my boyfriends like, 'Well what is this room? Where is she at on the list?'" [LS1]
- "He's still in the process of learning so when stuff happens, something like that, especially if we're going to a doctor... he can kind of answer the questions for me." [LS1]
- "Never go by yourself you really can't talk while you're in pain. You'll be sitting there." [LS4]
- "The mother is not going home. The mother will say, Doc, this is what the national guidelines say. Why you sending me home?" [LS5]

Physical care:

- "Yeah, they know, you know, I'm supposed to take folic acid... supposed to take a mild [non-steroidal anti-inflammatory] if I feel any minor pain coming on." [LS1]
- "Bring me some water or something or something." [LS3]
- "People try to help with the massages and all." [LS3]

Theme 6: Barriers to and Facilitators of Using an mHealth App

This theme emphasizes the perceived obstacles and benefits of using an mHealth app to assist with disease management (Textbox 6). We identified 3 levels of barriers and facilitators: patient, app-specific, and provider. Patient-level barriers included characteristics of the participants that would prevent them from using the application. This included apprehension toward internet use. For example, "I do community outreach and a lot of people are not computer savvy. Some people don't want to have nothing to do with the Internet" [LS3]. In addition, participants acknowledged that after consistent use of the application, there was the potential to ignore alerts:

That takes to not turning your notifications off your phone. Cause I'm quick to turn the notification off on everything cause that phone.... [LS4]

The facilitators explained the benefits of using an mHealth app. These factors would make it easier for the participants to use or become more interested in using the app. For example, a participant reported, "so I think it would be very helpful, especially for a sister or a spouse, whoever is with you that day" [LS4]. Another participant shared that a facilitator to using an app was as follows:

...if it had simple interface, something easily configurable that I don't have to go and do a lot...if it imports my medical records that would be fantastic. [LS1]

Textbox 6. Barriers to and facilitators of using a mobile health app. Illustrative quotes are examples from community listening participants recruited to develop the categories.

Illustrative quotes**Patient level:**

- “Doesn’t pay the phone bill.” [LS1]
- “The same thing that would prevent me from forgetting my medicine. I would get busy, the change in schedule.” [LS2]
- “Some people are just not familiar with an app.” [LS3]
- “If you can’t use internet, I don’t know what would make it easier.” [LS3]
- “The problem is, she has had strokes so she can’t do her hands in order to do things smartphone.” [LS3]
- “That takes to not turn your notifications off your phone. Cause I’m quick to turn the notification off on everything cause that phone get to bling, bling, bling, bling.” [LS4]
- “I probably open three of them because of practice and maybe one of them because of my own health.” [LS5]

App-specific:

- “If it’s not user-friendly you’re not gonna get much success out of it either thought.” [LS1]
- “Some apps ask you to connect to different servers with both monitoring and notification and stuff like that.” [LS2-context]
- “A technology, especially in the world that we’re in and some stuff that we’re kind of working on already, has the ability because everybody has smart phones.” [LS5]

Provider:

- “Are the doctors buying into this app, because it’s just an app?” [LS4]
- “Each doctor going to do something different.” [LS4]
- “Okay, hypothetically, you go into the emergency room. And you show them this app. This is what I need and this is my information. And he don’t follow it.” [LS4]

Theme 7: Feature Preferences for an mHealth App

Feature preferences included various features that would facilitate SCD self-management. Four key features were identified: alert system, communication with providers, information tracker, and the ability to interact with caregivers (Textbox 7). Participants stated that they wanted an alert system. For example, “Alerts you or reminds you to...Don’t forget your folic acid” [LS1]. Another participant stated, “Put it on the

phones so I can get a reminder” [LS4]. Participants wanted to be able to maintain communication with their providers through the mHealth app. One participant stated, “Even if the app was able to alert or connect to a patient registry or to allow them to email doctors if they are almost in a crisis” [LS2]. Other suggestions for an mHealth app were to have an information tracker and a feature that could enable an individual to interact with their caregiver.

Textbox 7. Feature preferences for a mobile health app. Illustrative quotes are examples from community listening participants recruited to develop the categories.

Illustrative quotes

Alert system:

- “Alerts you or reminds you to...Don't forget your folic acid.” [LS1]
- “I would actually like for it to give me alerts or triggers.” [LS1]
- “With me, when I'm puttin' in certain things like it tell you went over your carbs for the day or whatever put in the day it's something that children realize until you get the alert.” [LS2]
- “Put it on the phones so I can get a reminder.” [LS4]

Communication:

- “Even if the app was able to alert or log in somehow connect to a patient registry or whatever to allow them to email doctors to alert them if they are almost in a crisis based on the information.” [LS2]
- “I think the app should have a daily tracker. Just somebody came how's your mood today? How are you feeling today?” [LS2]
- “Is it like a way the app could be linked into the hospital or clinic's computer cause that way any info you putting in the app already will be there anyway. I mean, if that's the case then when you get to the emergency room or to the clinic... the computer already shows them whatever they need to know from your app.” [LS4]

Information tracker:

- “It shows like all your medications and the things you kind of track that stuff, track down the time that the nurses come in and give her medicine, what they giving her, how many milligrams.” [LS4]
- “I think the app should have a daily tracker. Just somebody came how's your mood today? How are you feeling today? Over time I think that information would be helpful when you do have to go see your doctor...” [LS2]

Ability to engage caregiver:

- “Just preventative things to that a caretaker can read.” [LS1]
- “You tell the caregiver that, you know, make sure you go... Give them time for the medicine to work kind of thing.” [LS1]
- “Giving some expected measures of time for a caregiver to say, you know, ‘Hey check on them. Let them rest. You know, no noise, stress relief, close the door, shut the blinds.’” [LS1]
- “You can have like a preventive section for the caregiver to also refer to before it even gets to the point where you're at the hospital.” [LS1]
- “Leave me alone. Go look at the app.” [LS3]
- “I was going to ask from a caregiver's perspective, we were talking about the app... an app where you kind of track that stuff, even stuff like, she's a hard stick so there might be one person in the whole hospital that can get a line off her. Something simple like recording the name of that person.” [LS4]

Discussion

Principal Findings

Studies have demonstrated the efficacy of mHealth self-management interventions in improving clinical health outcomes. This study describes the preferences for mHealth app features to facilitate self-management in adults with SCD and their caregivers residing in urban and rural communities. The patient and caregiver inputs confirmed key findings from the literature and exposed the gaps in realizing the needs of patients with SCD and their caregivers in disease management.

In our study, participants were interested in using an SCD mHealth app to increase their ability to communicate with health care providers. This was similar to the findings of other qualitative studies on the benefits of mHealth for patients with chronic health conditions [37,38]. Furthermore, studies suggest that the most effective mHealth interventions for disease self-management connect patients with their health care team

using 2-way communication [39,40]. Therefore, mHealth could be an effective method for meeting the patient communication needs. However, this communication was within the context of specialty care. One area that may not be as easily addressed through mHealth in SCD is communication in acute care settings.

The patients and caregivers in this study cited communication with providers in the emergency room as a challenge to disease management. In particular, they expressed concerns that they were not being heard and were often mistreated. These concerns are consistent with the health care experiences expressed among patients with SCD in other studies [41,42]. This reflects that although implementing mHealth in self-management interventions could potentially be effective in increasing communication with the health care providers, there is still a need to improve the SCD patient-provider interactions, particularly in acute settings.

Another major finding was the desire for patients to use mHealth to share SCD information with their caregivers. This was an expected finding, as participants also expressed that their main source of support for disease management was interpersonal support rather than a health care provider. Although few studies elicit preferences for mHealth among caregivers, obtaining feedback from them confirmed the benefit of adopting mHealth to facilitate communication between the patients and their caregivers [43,44]. Caregivers are often the main source of support for disease management in patients with chronic conditions, especially those residing in rural areas [45]. Health professionals or providers seeking to implement mHealth for expanding their services to rural patients should consider the usability of a proposed app for caregivers.

One consistent element in mHealth interventions is the use of reminder alerts or SMS text messaging [46,47]. Our study participants reported adhering to these alerts as a barrier to mHealth adoption. They stated that if they received several notifications from the app, they would shut off the notifications. This concern of alert fatigue, receiving a large quantity of information and having insufficient time or cognitive resources to distinguish relevant from irrelevant information [48,49], has also been described in provider-based health information technologies. As with this study's participants, the result was overriding or dismissing alerts [50,51]. Therefore, health care professionals and providers seeking to implement these elements into an mHealth intervention should consider their potential barriers before implementation. The suggested strategies for overcoming this barrier can include allowing the users to prioritize alerts and to specify the types of alerts they would prefer receiving [50,51].

Limitations

We acknowledge the limitations of this study. First, the community listening sessions included individuals with SCD

and their caregivers. It is possible that responses may have varied if these 2 groups participated in separate listening sessions. Next, as with other qualitative approaches, the results are not generalizable outside the sample population. Finally, sampling bias may have affected the findings, as participants who were willing to participate in the community listening sessions may differ from the general population of patients with SCD and their caregivers. In particular, the caregivers in our study were parents and spouses of adults with SCD. Their perspectives may differ from those of caregivers who are the parents of children and adolescents with SCD. Despite these limitations, we believe that the contributions of this study outweigh the limitations. As technology is increasingly incorporated into SCD management, it is important to capture the perspectives of the 2 important groups most impacted, which are the patients and their caregivers.

Conclusions

Patient and caregiver preferences are essential to understanding the components of mHealth technology that should be integrated into disease self-management to meet patient needs. This was one of the first studies to include the input from patients with SCD and their caregivers from rural communities. As technology-based strategies are continually developed to increase health care access in rural and medically underserved communities, it is important to incorporate the inputs from members of these communities. Future studies should assess the accessibility of technology resources available to patients with SCD residing in rural communities. These findings can be used to develop and enhance patient-centered, user-friendly mHealth apps that address the barriers to information access and facilitate disease self-management without creating an additional burden for the patients and their caregivers.

Acknowledgments

The authors would like to thank Vanessa Steele, Janice Reed, and the Sickle Cell Foundation of Tennessee for facilitating the community listening sessions. The authors dedicate this paper to Dr Trevor K Thompson, founder of the Sickle Cell Foundation of Tennessee. This program was funded, in part, through a Patient-Centered Outcomes Research Institute Eugene Washington Engagement Award (#1107) and the Vanderbilt University Trans-Institutional Projects program.

Conflicts of Interest

None declared.

References

1. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. *Am J Public Health* 2014 Aug;104(8):e25-e31. [doi: [10.2105/AJPH.2014.302041](https://doi.org/10.2105/AJPH.2014.302041)] [Medline: [24922170](https://pubmed.ncbi.nlm.nih.gov/24922170/)]
2. Wang T, Tan J, Xiao LD, Deng R. Effectiveness of disease-specific self-management education on health outcomes in patients with chronic obstructive pulmonary disease: an updated systematic review and meta-analysis. *Patient Educ Couns* 2017 Aug;100(8):1432-1446. [doi: [10.1016/j.pec.2017.02.026](https://doi.org/10.1016/j.pec.2017.02.026)] [Medline: [28318846](https://pubmed.ncbi.nlm.nih.gov/28318846/)]
3. Lepard MG, Joseph AL, Agne AA, Cherrington AL. Diabetes self-management interventions for adults with type 2 diabetes living in rural areas: a systematic literature review. *Curr Diab Rep* 2015 Jun;15(6):608 [FREE Full text] [doi: [10.1007/s11892-015-0608-3](https://doi.org/10.1007/s11892-015-0608-3)] [Medline: [25948497](https://pubmed.ncbi.nlm.nih.gov/25948497/)]

4. Gordon NP, Hornbrook MC. Differences in access to and preferences for using patient portals and other ehealth technologies based on race, ethnicity, and age: a database and survey study of seniors in a large health plan. *J Med Internet Res* 2016 Mar 04;18(3):e50 [FREE Full text] [doi: [10.2196/jmir.5105](https://doi.org/10.2196/jmir.5105)] [Medline: [26944212](https://pubmed.ncbi.nlm.nih.gov/26944212/)]
5. Christopher Gibbons M. Use of health information technology among racial and ethnic underserved communities. *Perspect Health Inf Manag* 2011 Jan 01;8:1f [FREE Full text] [Medline: [21307989](https://pubmed.ncbi.nlm.nih.gov/21307989/)]
6. Blumenthal D. Launching HITECH. *N Engl J Med* 2010 Feb 04;362(5):382-385. [doi: [10.1056/NEJMp0912825](https://doi.org/10.1056/NEJMp0912825)] [Medline: [20042745](https://pubmed.ncbi.nlm.nih.gov/20042745/)]
7. Lyles CR, Harris LT, Jordan L, Grothaus L, Wehnes L, Reid RJ, et al. Patient race/ethnicity and shared medical record use among diabetes patients. *Med Care* 2012 May 01;50(5):434-440. [doi: [10.1097/MLR.0b013e318249d81b](https://doi.org/10.1097/MLR.0b013e318249d81b)] [Medline: [22354209](https://pubmed.ncbi.nlm.nih.gov/22354209/)]
8. Sealy-Jefferson S, Vickers J, Elam A, Wilson MR. Racial and ethnic health disparities and the affordable care act: a status update. *J Racial Ethn Health Disparities* 2015 Dec;2(4):583-588 [FREE Full text] [doi: [10.1007/s40615-015-0113-z](https://doi.org/10.1007/s40615-015-0113-z)] [Medline: [26668787](https://pubmed.ncbi.nlm.nih.gov/26668787/)]
9. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. *J Med Internet Res* 2014 Jul 16;16(7):e172 [FREE Full text] [doi: [10.2196/jmir.3117](https://doi.org/10.2196/jmir.3117)] [Medline: [25048379](https://pubmed.ncbi.nlm.nih.gov/25048379/)]
10. Choi NG, Dinitto DM. Internet use among older adults: association with health needs, psychological capital, and social capital. *J Med Internet Res* 2013 May 16;15(5):e97 [FREE Full text] [doi: [10.2196/jmir.2333](https://doi.org/10.2196/jmir.2333)] [Medline: [23681083](https://pubmed.ncbi.nlm.nih.gov/23681083/)]
11. Lyles C, Schillinger D, Sarkar U. Connecting the dots: health information technology expansion and health disparities. *PLoS Med* 2015 Jul;12(7):e1001852 [FREE Full text] [doi: [10.1371/journal.pmed.1001852](https://doi.org/10.1371/journal.pmed.1001852)] [Medline: [26172977](https://pubmed.ncbi.nlm.nih.gov/26172977/)]
12. Lorence D, Park H. Study of education disparities and health information seeking behavior. *Cyberpsychol Behav* 2007 Feb;10(1):149-151. [doi: [10.1089/cpb.2006.9977](https://doi.org/10.1089/cpb.2006.9977)] [Medline: [17305464](https://pubmed.ncbi.nlm.nih.gov/17305464/)]
13. Sick cell disease. National Institutes of Health. URL: <https://www.nhlbi.nih.gov/health-topics/sickle-cell-disease> [accessed 2021-11-11]
14. Centers for Disease Control and Prevention. URL: <https://www.google.com/search?q=Sickle+cell+anemia%3A+who+is+at+risk%3F+.+In%3A+US+Department+of+Health+and+Human+Services> [accessed 2021-11-11]
15. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/ncbddd/sicklecell/data.html> [accessed 2021-11-11]
16. Lanzkron S, Carroll CP, Haywood C. Mortality rates and age at death from sickle cell disease: U.S., 1979-2005. *Public Health Rep* 2013;128(2):110-116 [FREE Full text] [doi: [10.1177/003335491312800206](https://doi.org/10.1177/003335491312800206)] [Medline: [23450875](https://pubmed.ncbi.nlm.nih.gov/23450875/)]
17. Ware RE, de Montalembert M, Tshilolo L, Abboud MR. Sickle cell disease. *Lancet* 2017 Jul 15;390(10091):311-323. [doi: [10.1016/S0140-6736\(17\)30193-9](https://doi.org/10.1016/S0140-6736(17)30193-9)] [Medline: [28159390](https://pubmed.ncbi.nlm.nih.gov/28159390/)]
18. Platt OS, Brambilla DJ, Rosse WF, Milner PF, Castro O, Steinberg MH, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med* 1994 Jun 09;330(23):1639-1644. [doi: [10.1056/NEJM199406093302303](https://doi.org/10.1056/NEJM199406093302303)] [Medline: [7993409](https://pubmed.ncbi.nlm.nih.gov/7993409/)]
19. Eltzschig HK, Eckle T. Ischemia and reperfusion--from mechanism to translation. *Nat Med* 2011 Nov 07;17(11):1391-1401 [FREE Full text] [doi: [10.1038/nm.2507](https://doi.org/10.1038/nm.2507)] [Medline: [22064429](https://pubmed.ncbi.nlm.nih.gov/22064429/)]
20. Azar S, Wong TE. Sickle cell disease: a brief update. *Med Clin North Am* 2017 Mar;101(2):375-393. [doi: [10.1016/j.mcna.2016.09.009](https://doi.org/10.1016/j.mcna.2016.09.009)] [Medline: [28189177](https://pubmed.ncbi.nlm.nih.gov/28189177/)]
21. Living well with sickle cell disease. National Center on Birth Defects and Developmental Disabilities. URL: https://www.cdc.gov/ncbddd/sicklecell/documents/livingwell-with-sickle-cell-disease_self-caretoolkit.pdf [accessed 2021-11-14]
22. The management of sickle cell disease. Division of Blood Diseases and Resources. URL: https://www.nhlbi.nih.gov/files/docs/guidelines/sc_mngt.pdf [accessed 2021-11-14]
23. Shah N, Jonassaint J, De Castro L. Patients welcome the Sickle Cell Disease Mobile Application to Record Symptoms via Technology (SMART). *Hemoglobin* 2014;38(2):99-103. [doi: [10.3109/03630269.2014.880716](https://doi.org/10.3109/03630269.2014.880716)] [Medline: [24512633](https://pubmed.ncbi.nlm.nih.gov/24512633/)]
24. Jonassaint C, Kang C, Prussien K, Yarboi J, Sanger MS, Wilson JD, et al. Feasibility of implementing mobile technology-delivered mental health treatment in routine adult sickle cell disease care. *Transl Behav Med* 2020 Feb 03;10(1):58-67 [FREE Full text] [doi: [10.1093/tbm/iby107](https://doi.org/10.1093/tbm/iby107)] [Medline: [30508141](https://pubmed.ncbi.nlm.nih.gov/30508141/)]
25. Crosby LE, Ware RE, Goldstein A, Walton A, Joffe NE, Vogel C, et al. Development and evaluation of iManage: a self-management app co-designed by adolescents with sickle cell disease. *Pediatr Blood Cancer* 2017 Jan;64(1):139-145 [FREE Full text] [doi: [10.1002/pbc.26177](https://doi.org/10.1002/pbc.26177)] [Medline: [27574031](https://pubmed.ncbi.nlm.nih.gov/27574031/)]
26. Badawy SM, Thompson AA, Liem RI. Technology access and smartphone app preferences for medication adherence in adolescents and young adults with sickle cell disease. *Pediatr Blood Cancer* 2016 May 04;63(5):848-852. [doi: [10.1002/pbc.25905](https://doi.org/10.1002/pbc.25905)] [Medline: [26844685](https://pubmed.ncbi.nlm.nih.gov/26844685/)]
27. Matthie N, Hamilton J, Wells D, Jenerette C. Perceptions of young adults with sickle cell disease concerning their disease experience. *J Adv Nurs* 2016 Jun;72(6):1441-1451 [FREE Full text] [doi: [10.1111/jan.12760](https://doi.org/10.1111/jan.12760)] [Medline: [26350494](https://pubmed.ncbi.nlm.nih.gov/26350494/)]
28. Lee E, Pieczynski J, DeDios-Stern S, Simonetti C, Lee GK. Gender differences in caregiver strain, needs for support, social support, and quality of life among spousal caregivers of persons with multiple sclerosis. *Work* 2015;52(4):777-787. [doi: [10.3233/WOR-152205](https://doi.org/10.3233/WOR-152205)] [Medline: [26599674](https://pubmed.ncbi.nlm.nih.gov/26599674/)]

29. Geng HM, Chuang DM, Yang F, Yang Y, Liu WM, Liu LH, et al. Prevalence and determinants of depression in caregivers of cancer patients: a systematic review and meta-analysis. *Medicine (Baltimore)* 2018 Sep;97(39):e11863-e11870 [FREE Full text] [doi: [10.1097/MD.00000000000011863](https://doi.org/10.1097/MD.00000000000011863)] [Medline: [30278483](https://pubmed.ncbi.nlm.nih.gov/30278483/)]
30. Gordon P, Perrone K. When spouses become caregivers: counseling implications for younger couples. *J Rehab* 2004 Apr 01;70(2):27-32 [FREE Full text] [doi: [10.4135/9781483369532.n59](https://doi.org/10.4135/9781483369532.n59)]
31. Roadmap to livability collection: community listening session toolkit. AARP. URL: <https://www.aarp.org/content/dam/aarp/livable-communities/livable-documents/documents-2018/Book-2-Roadmap-To-Livability-Listening-21318-lr.pdf> [accessed 2021-11-14]
32. Erves JC, Mayo-Gamble TL, Malin-Fair A, Boyer A, Joosten Y, Vaughn YC, et al. Needs, priorities, and recommendations for engaging underrepresented populations in clinical research: a community perspective. *J Community Health* 2017 Jun;42(3):472-480 [FREE Full text] [doi: [10.1007/s10900-016-0279-2](https://doi.org/10.1007/s10900-016-0279-2)] [Medline: [27812847](https://pubmed.ncbi.nlm.nih.gov/27812847/)]
33. Mayo-Gamble TL, Murry VM, Cunningham-Erves J, Cronin RM, Lari N, Gorden A, et al. Engaging individuals with sickle cell disease in patient-centered outcomes research: a community health ambassador training model. *J Health Care Poor Underserved* 2020;31(1):353-369. [doi: [10.1353/hpu.2020.0027](https://doi.org/10.1353/hpu.2020.0027)] [Medline: [32037336](https://pubmed.ncbi.nlm.nih.gov/32037336/)]
34. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods* 2016 Nov 29;5(1):80-92. [doi: [10.1177/160940690600500107](https://doi.org/10.1177/160940690600500107)]
35. 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/)]
36. Geisler C, Swarts J. *Coding Streams of Language: Techniques for the Systematic Coding of Text, Talk, and Other Verbal Data Streams of Language*. Boulder, CO: University Press of Colorado; 2020.
37. O'Connor S. Co-designing technology with people with dementia and their carers: exploring user perspectives when co-creating a mobile health application. *Int J Older People Nurs* 2020 Sep 14;15(3):e12288-e12297. [doi: [10.1111/opn.12288](https://doi.org/10.1111/opn.12288)] [Medline: [31837096](https://pubmed.ncbi.nlm.nih.gov/31837096/)]
38. Korpershoek YJ, Vervoort SC, Trappenburg JC, Schuurmans MJ. Perceptions of patients with chronic obstructive pulmonary disease and their health care providers towards using mHealth for self-management of exacerbations: a qualitative study. *BMC Health Serv Res* 2018 Oct 04;18(1):757-770 [FREE Full text] [doi: [10.1186/s12913-018-3545-4](https://doi.org/10.1186/s12913-018-3545-4)] [Medline: [30286761](https://pubmed.ncbi.nlm.nih.gov/30286761/)]
39. Greenwood DA, Gee PM, Fatkin KJ, Peoples 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/)]
40. Farzandipour M, Nabovati E, Sharif R, Arani M, Anvari S. Patient self-management of asthma using mobile health applications: a systematic review of the functionalities and effects. *Appl Clin Inform* 2017 Dec 14;08(04):1068-1081. [doi: [10.4338/aci-2017-07-r-0116](https://doi.org/10.4338/aci-2017-07-r-0116)] [Medline: [29241254](https://pubmed.ncbi.nlm.nih.gov/29241254/)]
41. Renedo A, Miles S, Chakravorty S, Leigh A, Telfer P, Warner JO, et al. Not being heard: barriers to high quality unplanned hospital care during young people's transition to adult services - evidence from 'this sickle cell life' research. *BMC Health Serv Res* 2019 Nov 21;19(1):876-887 [FREE Full text] [doi: [10.1186/s12913-019-4726-5](https://doi.org/10.1186/s12913-019-4726-5)] [Medline: [31752858](https://pubmed.ncbi.nlm.nih.gov/31752858/)]
42. Dyal BW, Abudawood K, Schoppee TM, Jean S, Smith VM, Greenlee A, et al. Reflections of healthcare experiences of African Americans with sickle cell disease or cancer: a qualitative study. *Cancer Nurs* 2021;44(1):E53-E61. [doi: [10.1097/NCC.0000000000000750](https://doi.org/10.1097/NCC.0000000000000750)] [Medline: [31743153](https://pubmed.ncbi.nlm.nih.gov/31743153/)]
43. Raval MV, Taylor N, Piper K, Thakore M, Hoff K, Owens S, et al. Pediatric patient and caregiver preferences in the development of a mobile health application for management of surgical colorectal conditions. *J Med Syst* 2017 Jul;41(7):105. [doi: [10.1007/s10916-017-0750-3](https://doi.org/10.1007/s10916-017-0750-3)] [Medline: [28540617](https://pubmed.ncbi.nlm.nih.gov/28540617/)]
44. Xie B, Champion JD, Kwak J, Fleischmann KR. Mobile health, information preferences, and surrogate decision-making preferences of family caregivers of people with dementia in rural hispanic communities: cross-sectional questionnaire study. *J Med Internet Res* 2018 Dec 10;20(12):e11682 [FREE Full text] [doi: [10.2196/11682](https://doi.org/10.2196/11682)] [Medline: [30530450](https://pubmed.ncbi.nlm.nih.gov/30530450/)]
45. Angwenyi V, Aantjes C, Kajumi M, De Man J, Criel B, Bunders-Aelen J. Patients experiences of self-management and strategies for dealing with chronic conditions in rural Malawi. *PLoS One* 2018 Jul 2;13(7):e0199977 [FREE Full text] [doi: [10.1371/journal.pone.0199977](https://doi.org/10.1371/journal.pone.0199977)] [Medline: [29965990](https://pubmed.ncbi.nlm.nih.gov/29965990/)]
46. 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/)]
47. 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/)]
48. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006;13(2):138-147 [FREE Full text] [doi: [10.1197/jamia.M1809](https://doi.org/10.1197/jamia.M1809)] [Medline: [16357358](https://pubmed.ncbi.nlm.nih.gov/16357358/)]
49. van der Sijs H, van Gelder T, Vulto A, Berg M, Aarts J. Understanding handling of drug safety alerts: a simulation study. *Int J Med Inform* 2010 May;79(5):361-369. [doi: [10.1016/j.ijmedinf.2010.01.008](https://doi.org/10.1016/j.ijmedinf.2010.01.008)] [Medline: [20171929](https://pubmed.ncbi.nlm.nih.gov/20171929/)]

50. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, with the HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* 2017 Apr 10;17(1):36 [FREE Full text] [doi: [10.1186/s12911-017-0430-8](https://doi.org/10.1186/s12911-017-0430-8)] [Medline: [28395667](https://pubmed.ncbi.nlm.nih.gov/28395667/)]
51. Kane-Gill SL, O'Connor MF, Rothschild JM, Selby NM, McLean B, Bonafide CP, et al. Technologic distractions (part 1): summary of approaches to manage alert quantity with intent to reduce alert fatigue and suggestions for alert fatigue metrics. *Crit Care Med* 2017 Sep;45(9):1481-1488. [doi: [10.1097/CCM.0000000000002580](https://doi.org/10.1097/CCM.0000000000002580)] [Medline: [28682835](https://pubmed.ncbi.nlm.nih.gov/28682835/)]

Abbreviations

mHealth: mobile health

SCD: sickle cell disease

SCFT: Sickle Cell Foundation of Tennessee

Edited by G Eysenbach; submitted 10.03.21; peer-reviewed by S Creary, K Fuji, A Hidki; comments to author 12.05.21; revised version received 26.07.21; accepted 19.09.21; published 30.11.21.

Please cite as:

*Mayo-Gamble TL, Quasie-Woode D, Cunningham-Erves J, Rollins M, Schlundt D, Bonnet K, Murry VM
Preferences for Using a Mobile App in Sickle Cell Disease Self-management: Descriptive Qualitative Study
JMIR Form Res 2021;5(11):e28678*

URL: <https://formative.jmir.org/2021/11/e28678>

doi: [10.2196/28678](https://doi.org/10.2196/28678)

PMID: [34851295](https://pubmed.ncbi.nlm.nih.gov/34851295/)

©Tilicia L Mayo-Gamble, Delores Quasie-Woode, Jennifer Cunningham-Erves, Margo Rollins, David Schlundt, Kemberlee Bonnet, Velma McBride Murry. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 30.11.2021. 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

Body Image Distress and Its Associations From an International Sample of Men and Women Across the Adult Life Span: Web-Based Survey Study

Alyssa Milton^{1,2*}, BSc, MAppSc, PhD; Ashlea Hambleton^{3,4*}, BPsych, MCLinPsych; Anna Roberts^{1,3}, BAppSc, MExSc, MBMSc; Tracey Davenport^{1,3}, BA, eMBA; Anna Flego⁵, BPT, MPH; Jane Burns³, PhD; Ian Hickie¹, AM, MD

¹Brain and Mind Centre, The University of Sydney, Sydney, Australia

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

³Project Synergy, InnoWell Pty Ltd, Sydney, Australia

⁴Inside Out Institute for Eating Disorders, Charles Perkins Centre, University of Sydney, Sydney, Australia

⁵The Movember Foundation, Australia, East Melbourne, Australia

* these authors contributed equally

Corresponding Author:

Alyssa Milton, BSc, MAppSc, PhD

Sydney School of Medicine

Faculty of Medicine and Health

University of Sydney

Professor Marie Bashir Centre

67-73 Missenden Rd

Camperdown, 2050

Australia

Phone: 61 02 93510774

Email: alyssa.milton@sydney.edu.au

Abstract

Background: Previous research on body image distress mainly relied on samples that were small, generally homogeneous in age or sex, often limited to one geographical region, and were characterized by a lack of comprehensive analysis of multiple psychosocial domains. The research presented in this paper extends the international literature using the results of the web-based Global Health and Wellbeing Survey 2015. The survey included a large sample of both men and women aged ≥ 16 years from Australia, Canada, New Zealand, the United Kingdom, or the United States.

Objective: The main objectives of this study are to examine body image distress across the adult life span (≥ 16 years) and sex and assess the association between body image distress and various psychosocial risk and protective factors.

Methods: Data were extracted from the Global Health and Wellbeing Survey 2015, a web-based international self-report survey with 10,765 respondents, and compared with previous web-based surveys conducted in 2009 and 2012.

Results: The body image distress of young Australians (aged 16-25 years) significantly rose by 33% from 2009 to 2015. In 2015, 75.19% (961/1278) of 16- to 25-year-old adults reported body image distress worldwide, and a decline in body image distress was noted with increasing age. More women reported higher levels of body image distress than men (1953/3338, 58.51% vs 853/2175, 39.22%). Sex, age, current dieting status, perception of weight, psychological distress, alcohol and other substance misuse, and well-being significantly explained 24% of the variance in body image distress in a linear regression ($F_{15,4966}=105.8$; $P<.001$).

Conclusions: This study demonstrates the significant interplay between body image distress and psychosocial factors across age and sex.

(JMIR Form Res 2021;5(11):e25329) doi:[10.2196/25329](https://doi.org/10.2196/25329)

KEYWORDS

body image; mental health; well-being; web-based survey; sex differences; age

Introduction

Background

There is a clear need to understand and address body image distress, particularly when considering the increasing prevalence rates of body image distress worldwide [1-3] and the noted relationship between body image distress and mental ill-health [4]. Furthermore, there is scant large-scale international research examining body image across the adult life span [4] and from the perspectives of both men and women [5].

Body image is a multifaceted construct encompassing one's body-related self-perceptions and self-attitudes, including thoughts, feelings, behaviors, and beliefs toward the body [6]. Research suggests that body image dissatisfaction occurs when there is a discrepancy between how an individual views their body (*actual body image*) and how they want it to be (*ideal body image*) [7,8]. Dissatisfaction, overevaluation, and preoccupation are considered as contributing factors to body image distress [9]. Current body ideals predominantly promote thinness for women and muscularity for men [10,11]. Given the difference between current body ideals and the actual body shape and size of most of the population, it is not surprising that many people view their bodies negatively and experience distress because of this negative self-perception [12].

Body Image Distress Over Time

Body image is ranked as a top concern for young people [13]. Research suggests that the proportion of the population experiencing increased body image distress is increasing. It has been well documented that viewing appearance-focused media contributes to the development of body image concerns [14,15]. In the past 10 years, there has been a stark rise in multimedia platforms, such as Instagram, Snapchat, and TikTok. The imagery on social media is often filtered and edited in a way that promotes an unrealistic appearance ideal. Research has shown that social media use is associated with increased body image dissatisfaction [14] because it facilitates comparison, and appearance-related comments and praise are reinforced with likes, follows, and comments.

Body Image, Age, and Sex

The association between weight status (as measured by BMI) and body image dissatisfaction has been previously explored [16]. Research has demonstrated that increased BMI is associated with greater body image dissatisfaction in college students [17], help-seeking adults [18], adolescent men [19], and nonclinical samples of adult women [20]. Risk of body dissatisfaction is not restricted to individuals with higher BMI; adolescent women with either a healthy or an overweight BMI experience higher levels of body dissatisfaction, whereas underweight women have the highest levels of body satisfaction [19]. Similarly, in a sample of middle-aged women, 70% of participants reported a desire to be thinner, despite most being considered normal weight [21]. Overall, these findings fit the current cultural narrative that a thin body is both desirable and idealized in Western societies [22].

Although the research has mainly focused on young women, some studies indicate that body image concerns are pervasive

across the adult life span [12] for both sexes [5]. The picture appears to be more complex for men [23]. For example, adolescent men tend to be equally divided between wanting to lose weight (predominantly high body fat) and gain weight (muscle mass) [5,24]. As men move into adulthood, there is an increase in the desire to lose weight [5]. Within cohorts of middle-aged women, only 11% of participants endorsed being satisfied with their bodies [25]. Women's dissatisfaction with their bodies appears relatively stable across the adult life span [8,12]. However, some research suggests that the impact of body image on an individual's self-esteem and self-concept may diminish over time [26]. Furthermore, there is more tolerance in what body sizes are considered acceptable with increasing age [4,27]. However, overall, body image research looking at age, sex, and weight is fragmented, and a comprehensive picture is lacking.

Body Image Distress and Psychopathology

Body image dissatisfaction in childhood and early adolescence can predict adverse health outcomes in later life, including engaging in dangerous weight control behaviors and general psychological distress [28]. As highlighted above, research with adolescents is much more extensive than with their adult counterparts, with several studies demonstrating an association between body dissatisfaction and anxiety [29-33], depression [29,30,34-37], self-harm [38-42], and low self-esteem [28,36,37].

Research has reported that body image dissatisfaction is associated with higher levels of depression, anxiety, disordered eating [43-45], and distress [46]. Furthermore, research has identified associations between body image with other aspects of health, such as tobacco smoking [47], alcohol misuse [48], poor self-esteem [18,49,50], and poor mental and physical health-related quality of life [46].

Conversely, optimism, positive affect, self-compassion, life satisfaction, and subjective happiness [51-61] are associated with positive body image. Social well-being has also been reported to play a part in both positive and negative body image, particularly in adolescents [59,62-64]. For example, Bearman et al [63] observed higher levels of body dissatisfaction in girls and boys who had deficits in their social support from parents and peers. Meanwhile, individuals with more supportive parental relationships have reported higher body image satisfaction [62].

Current Research

Previous research exploring body image distress mainly relied on small samples that were generally homogeneous in age or sex, lacked a comprehensive analysis of multiple psychosocial domains, and were limited to one geographical region. This study extends the international literature using the results of the Global Health and Wellbeing Survey 2015, a large web-based sample of both men and women (aged ≥ 16 years) from Australia, Canada, New Zealand, the United Kingdom, or the United States. Additional data sources included the *headspace* web-based Community Youth Survey (2009) and the Young and Well First National Survey on the web (2012).

This study has 3 main aims, including the assessment of (1) the changes in body image distress over time (between 2009 and

2015) for young people aged 16 to 25 years; (2) the associations between weight range (BMI), dieting status, and perceived body image distress by sex and age group; and (3) the association of various demographic, health, and well-being factors with body image distress or preoccupation.

Methods

Participants

Participants were a voluntary community sample of men and women (aged ≥ 16 years) who reported that they had lived in 1 of the 5 target countries (Australia, Canada, New Zealand, the United Kingdom, or the United States) for the best part of the past 12 months. A total of 16,510 people reviewed the consent and eligibility screen. Of the 16,510 people, the total eligible sample was 10,765 (65.2%) respondents. Of those excluded, 26.2% (4326/16,510) did not consent to participate, and 4.3% (710/16,510) were younger than 16 years.

Procedures and Recruitment

The primary study received institutional ethics approval from The University of Sydney Human Research Ethics Committee (protocol 2015/412). All procedures complied with the ethical standards of the relevant national and institutional committees on human ethics and the Helsinki Declaration of 1975, as revised in 2008.

The survey was hosted on the internet from July 1, 2015, to December 11, 2015. For optimizing recruitment in the 5 target countries, the following strategies were used: both paid and free advertising across multiple social media channels such as Facebook, Twitter, and YouTube for survey dissemination [65]; layering of recruitment messages [66]; passive web-based snowballing via social media to spread study information through sharing, liking, and tweeting [65,67]; and traditional snowballing [68]. Targeted recruitment based on age, sex, and region was carried out through paid advertising channels to maximize responses from groups hard to reach. Respondents consented on the website and were informed that their responses were confidential, nonidentifiable, and that they could cease participation at any time. Participants were informed that the survey would take between 20 and 45 minutes to complete depending on participant answers and the survey skip pattern. Any respondents indicating psychological distress or problematic alcohol or substance use in their responses were provided with the contact details of local support lines.

Materials

Items in this substudy were extracted from the Global Health and Wellbeing Survey 2015 [69], and where items could be compared, from the *headspace* web-based Community Youth Survey (2009) and the Young and Well First National Survey on the web (2012). Areas of interest specific to this substudy are described in the following sections.

Demographics

The respondents first provided their sex (men and women) and age (16-25 years, 26-49 years, or ≥ 50 years).

Weight, Body Image, and Eating Behaviors

BMI was determined by asking respondents their weight (kg or lb) and height (meters or feet and inches). Respondents with a BMI less than 18.5 were classified as *underweight*, a BMI between 18.5 and 24.9 as *healthy weight*, a BMI between 25.0 and 29.9 as *overweight*, and a BMI higher than 30.0 as *obese*.

Current dieting status was determined by the question "Are you currently dieting?" adapted from Blashill and Wilhelm [70]; this item provided 3 response options ("yes, to lose weight"; "yes, to gain weight"; and "no").

For assessing body image attitudes, respondents were asked to self-evaluate the importance of weight and shape for them over the past 3 months: "How much has your shape influenced how you feel as a person?" This question was answered on a 6-point Likert scale ranging from *not at all* to *a great deal*.

For assessing body image distress or preoccupation, respondents were asked, "Do you get very distressed or preoccupied by any specific aspect of your physical appearance?" using a dichotomous *yes* or *no* response option. If respondents indicated distress or preoccupation, a follow-up question asked which areas of their body they were concerned about, such as *facial features*, *arms or legs*, and *weight* [71]. This body image distress item was also asked in the web-based *headspace* Community Youth Survey in 2009 and the Young and Well National Survey on the web in 2012 with 16- to 25-year-old adults [71]. These data were used in this research for longitudinal cohort comparisons.

Mental Health and Well-being

Physical activity was measured by the International Physical Activity Questionnaire short form, which classifies individuals into 1 of 3 levels of physical activity (*inactive* vs *minimally active* vs *health-enhancing physical activity*) [72].

Current psychological distress was measured using the 10-item Kessler Psychological Distress Scale (K10) [73]. Total scores were grouped into 4 levels of psychological distress (10-15=low, 16-21=moderate, 22-29=high, and 30-50=very high) [74].

Respondents' levels of suicidal thoughts and behaviors in the past 12 months were measured using the 5-item suicidal thoughts and acts subscale from the Psychiatric Symptom Frequency Scale [75].

The likelihood of alcohol or other substance misuse was calculated using 2 items. If respondents positively endorsed one of either item, "...recently thought that you should cut down..." or "...another person suggested you should cut down..." they were categorized as having a *possible* alcohol or other substance misuse. Endorsement of both items resulted in *probable* alcohol or other substance misuse. Endorsement of neither item resulted in being placed in the *not likely* category.

Days out of role was extracted from the Brief Disability Questionnaire [76] to investigate functioning. The 7-item Personal Well-being Index [77] was used to assess subjective well-being.

Happiness was measured by the 4-item Oxford Happiness Questionnaire [78], and resilience was measured by the 4-item Brief Resilience Coping Scale [79].

Perceived social support and conflict in close relationships were measured by the 5-item Schuster Social Support and Conflict Scale [80].

Analysis

Survey data were prepared and analyzed using IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp, 2013). For addressing aim 1, body image distress data in 2009 and 2012 were compared with 2015 data using a one-sample two-tailed *t* test. To address aim 2, descriptive and frequency statistics were used to describe all weight, dieting, and body image items by basic demographics (sex and age). Chi-square and analysis of variance tests were completed to compare items by sex and age. For addressing aim 3, an initial scoping analysis using Pearson product-moment *r* correlations [81] was conducted to independently assess the strength of the relationship between body image distress and these health and well-being items (healthy weight; *no* vs *yes* based on BMI data), current dieting status (*no* vs *yes*), perception of weight (*about the right weight*

vs *all others*), physical activity (International Physical Activity Questionnaire short form), psychological distress (K10), suicidal ideation (Psychiatric Symptom Frequency Scale), alcohol and/or other substance misuse, days out of role, well-being (Personal Well-being Index), happiness (Oxford Happiness Questionnaire), resilience (Brief Resilience Coping Scale), intimate bonds (Intimate Bond Measure), and social support (Social Support and Conflict Scale). A subsequent linear regression analysis was conducted to determine which of the health and well-being items significantly explained variance in body image distress when considered together while controlling for sex and age.

Results

Respondent Participation Rates and Characteristics

Most of the eligible respondents were women (6464/10,765, 60.05%). Of the 10,753 respondents, 2874 (26.73%) were aged between 16 and 25 years, 2879 (26.77%) were aged between 26 and 49 years, and 5000 (46.49%) were aged ≥ 50 years. A breakdown of demographics by country, age, and sex is presented in Table 1. Further demographic details can be found in related publications [82,83].

Table 1. Participant demographics by country, age, and sex (N=10,765).

Characteristics	Total	Australia	Canada	New Zealand	United Kingdom	United States
Country, n (%)	10,765 (100)	3349 (31.11)	1888 (17.54)	1752 (16.27)	1938 (18)	1838 (17.07)
Sex, n (%)						
Women	6464 (60.05)	2067 (61.72)	1097 (58.1)	1055 (60.22)	1176 (60.68)	1069 (58.16)
Men	4301 (39.95)	1282 (38.28)	791 (41.89)	697 (39.78)	762 (39.32)	769 (41.84)
Age (years), mean (SD)	44.37 (19.68)	42.44 (18.71)	52.46 (17.84)	39.65 (19.51)	42.57 (19.40)	45.97 (21.07)

Main Findings

Aim 1: Changes in Body Image Distress Over Time Across the Life Span

Of those young Australians (aged 16-25 years) who completed the 2009 (*headspace* web-based Community Youth Survey), 2012 (Young and Well First National Survey on the web), or 2015 (Global Health and Wellbeing Survey) surveys, self-reported body image distress rose by 33% from 2009 to 2015 (419/949, 44.2% reported distress in 2009; 1158/1731, 66.89% in 2012; and 300/388, 77.32% in 2015). The mean difference was significant across both the 2009 and 2012 time points (2009 vs 2015: $t_{387}=15.56$, $P<.001$; 2012 vs 2015: $t_{387}=4.90$, $P<.001$).

Aim 2: Body Image, BMI, and Dieting by Age and Sex

Table 2 displays frequency statistics for measures of body image items, BMI, and current dieting status by age band and sex from the Global Health and Wellbeing Survey 2015. Significance tests comparing items across sex and age are presented in Multimedia Appendix 1. Approximately half of all respondents (2806/5513, 50.89%) reported feeling very distressed or preoccupied with their body image. Women reported higher levels of distress related to body image than men (1953/3338, 58.51% vs 853/2175, 39.22%). Respondents aged 16-25 years showed higher levels of body image distress than all other age groups, at 75.19% (961/1278), with distress decreasing as age increased.

Table 2. Frequency statistics for measures of body image item by age and sex (maximum N=5517).

Body image item	Sex		Age bands (years)			Full sample
	Men	Women	16-25	26-49	≥50	
Body image distress or preoccupation, n (%)	2175 (39.45)	3338 (60.55)	1278 (23.18)	1502 (27.24)	2733 (49.57)	5513 (100)
Yes	853 (39.22)	1953 (58.51)	961 (75.19)	837 (55.73)	1008 (36.88)	2806 (50.89)
No	1322 (60.78)	1385 (41.49)	317 (24.8)	665 (44.27)	1725 (63.12)	2707 (49.1)
How much does weight or shape influence how you think of yourself as a person	2169 (39.47)	3327 (60.53)	1275 (23.19)	1497 (27.24)	2724 (49.56)	5496 (100)
1=not at all, n (%)	510 (23.51)	425 (12.77)	130 (10.17)	181 (12.09)	624 (22.91)	935 (17.01)
6=a great deal, n (%)	256 (11.8)	771 (23.17)	330 (25.88)	321 (21.44)	376 (13.8)	1027 (18.69)
Score, median (IQR)	3 (2-4)	4 (2-5)	4 (3-6)	4 (2-5)	3 (2-5)	4 (2-5)
BMI, n (%)	2115 (39.32)	3264 (60.68)	1248 (23.2)	1464 (27.22)	2667 (49.58)	5379 (100)
Underweight	30 (1.42)	141 (4.32)	97 (7.77)	39 (2.67)	35 (1.31)	171 (3.18)
Healthy weight	671 (31.73)	1344 (41.19)	721 (57.77)	580 (39.64)	714 (26.77)	2015 (37.47)
Overweight	791 (37.39)	836 (25.62)	232 (18.59)	438 (29.94)	957 (35.88)	1627 (30.25)
Obese	623 (29.46)	842 (28.87)	198 (15.87)	406 (27.75)	961 (36.03)	1565 (29.09)
Self-evaluation of weight, n (%)	2176 (39.46)	3339 (60.54)	1279 (23.19)	1503 (27.25)	2733 (49.56)	5515 (100)
Very underweight	25 (1.15)	19 (0.56)	16 (1.25)	9 (0.59)	19 (0.69)	44 (0.79)
Slightly underweight	154 (7.08)	143 (4.28)	129 (10.09)	74 (4.92)	94 (3.44)	297 (5.39)
About the right weight	646 (29.69)	1057 (31.66)	574 (44.88)	522 (34.73)	607 (22.21)	1703 (30.88)
Slightly overweight	1000 (45.96)	1348 (40.37)	429 (33.54)	624 (41.52)	1295 (47.38)	2348 (42.57)
Very overweight	351 (16.13)	772 (23.12)	131 (10.24)	274 (18.23)	718 (26.27)	1123 (20.36)
Currently dieting, n (%)	2178 (39.48)	3339 (60.52)	1278 (23.16)	1503 (27.24)	2736 (49.59)	5517 (100)
Yes, to lose weight	427 (19.61)	868 (25.99)	268 (20.97)	374 (24.88)	653 (23.87)	1295 (23.47)
Yes, to gain weight	49 (2.25)	27 (0.81)	38 (2.97)	17 (1.13)	21 (0.77)	76 (1.38)
No	1702 (78.15)	2444 (73.20)	972 (76.06)	1112 (73.99)	2062 (75.37)	4146 (75.15)

Only 17.01% (935/5496) of respondents in the full sample indicated that their weight and shape did not influence how they thought of themselves as a person. Women (425/3327, 12.77%) and the younger age bands (16-25 years: 130/1275, 10.19%; 25-44 years: 181/1497, 12.09%) were significantly less likely to endorse that their weight or shape did not influence their self-perception ($P<.001$).

Although 57.77% (721/1248) of young people (16-25 years) were in the healthy BMI range, fewer (574/1279, 44.88%) considered themselves about the right weight. This pattern was repeated in women, of whom 41.19% (1344/3264) were in the healthy BMI range, but fewer (1057/3339, 31.66%) endorsed that they were about the right weight. The percentage of men who were in a healthy BMI range was 31.73% (671/2115), which reflected the rates of men who felt they were about the right weight (646/2176, 29.69%). In the older age brackets, more participants were in the obese BMI category (26-49 years: 406/1464, 27.75%; ≥50 years: 961/2667, 36.03%) than those who felt that they were very overweight (26-49 years: 274/1503, 18.23%; ≥50 years: 718/2733, 26.27%). Across the full sample, 23.47% (1295/5517) of all participants were currently dieting

to lose weight, and 1.38% (76/5517) were currently dieting to gain weight. Across ages, not engaging in any dieting was relatively consistent (between 1112/1503, 73.99% and 972/1278, 76.06%). Women reported they were dieting to lose weight significantly more frequently (868/3339, 25.99%) than men (427/2178, 19.61%; $P<.001$).

For addressing aim 3, a series of linear regressions were conducted to examine the relationship between body image distress or preoccupation and health and well-being items. [Table 3](#) presents the full sample results ([Multimedia Appendix 2](#) for each age group). Individual Pearson product-moment r correlations for each health and well-being item by body image distress are presented in [Multimedia Appendix 3](#). The regression model using the full sample significantly accounted for 24% of the variance in body image distress or preoccupation ($F_{15,4966}=105.8$; $P<.001$; $R^2_{adj}=0.24$). After controlling for sex ($\beta=.12$; $P<.001$) and age ($\beta=-0.24$; $P<.001$), 5 variables significantly explained model variance. This included current dieting status ($\beta=.13$; $P<.001$), perception of weight ($\beta=.09$; $P<.001$), psychological distress ($\beta=.21$; $P<.001$), alcohol and/or other substance misuse ($\beta=.04$; $P<.001$), and well-being

($\beta=-0.07$; $P<.001$). Specifically, respondents who were currently dieting reported body image distress or preoccupation more frequently. Those who did not report that they were *about the right weight* reported higher psychological distress and had a

higher likelihood of problematic alcohol or other substance use and higher body image distress or preoccupation. Participants with higher personal well-being scores reported lower levels of body image distress or preoccupation.

Table 3. Linear regression of body image distress ($F_{15,4966}=105.8$; $P<.001$; $R^2_{adj}=0.24$).

Variable	t (df) ^a	P value	β (95% CI)
Body image distress or preoccupation			
Healthy weight (no vs yes based on BMI)	0.04	.97	.00 (–0.03 to 0.03)
Current dieting (no vs yes)	10.20	<.001	.13 (0.12 to 0.18)
Perception of weight (about the right weight vs not)	6.06	<.001	.09 (0.07 to 0.13)
Physical activity (IPAQ ^b)	–0.03	.97	.00 (–0.02 to 0.02)
Psychological distress (K10 ^c)	9.45	<.001	.21 (0.01 to 0.01)
Suicidal ideation (PSFS ^d)	0.38	.71	.01 (–0.03 to 0.04)
Alcohol and/or other substance misuse	2.82	<.001	.04 (0.01 to 0.04)
Days out of role	–1.42	.16	–0.02 (–0.01 to 0.00)
Well-being (PWI ^e)	–3.37	<.001	–0.07 (0.00 to 0.00)
Happiness (OHQ ^f)	–0.40	.69	–0.01 (–0.01 to 0.00)
Resilience (BRCS ^g)	0.84	.40	.01 (0.00 to 0.01)
Social support (SSCS ^h)	–1.39	.16	–0.02 (–0.01 to 0.00)
Intimate bonds (IBM ⁱ)	–0.57	.57	–0.01 (0.00 to 0.00)
Sex	9.46	<.001	.12 (0.10 to 0.15)
Age	–16.57	<.001	–0.24 (–0.01 to –0.01)

^a $df=15,4966$

^bIPAQ: International Physical Activity Questionnaire short form.

^cK10: 10-item Kessler Psychological Distress Scale.

^dPSFS: Psychiatric Symptom Frequency Scale.

^ePWI: Personal Well-being Index.

^fOHQ: Oxford Happiness Questionnaire.

^gBRCS: Brief Resilience Coping Scale.

^hSSCS: Schuster Social Support and Conflict Scale.

ⁱIBM: Intimate Bond Measure.

When analyzed separately by age group ([Multimedia Appendix 2](#)), sex, current dieting status, perception of weight, and psychological distress consistently explained model variance across all age groups. Variation was found for happiness, alcohol or other substance misuse, and well-being items. Specifically, lower happiness also explained body image distress ($\beta=-0.16$; $P=.003$) in young people (aged 16–25 years). For those aged 26 to 49 years, alcohol and/or other substance misuse remained an item that explained body image distress ($\beta=.07$; $P=.008$). Conversely, for the ≥ 50 years age group, lower well-being continued to explain body image distress ($\beta=-0.10$; $P=.001$) variance.

Discussion

Principal Findings

To our knowledge, this is the largest international study to examine body image distress—and other related factors, including self-reported and perceived weight range and dieting status—across time, age, and sex. Our findings show that body image distress has become a highly prevalent issue by 2015. Of concern, considerable levels of body image distress were present in women and young people, and multiple psychosocial risk factors were associated with this distress.

One of our key findings comes from the cross-sectional longitudinal Australian data. Self-reported body image distress in young people aged 16 to 25 years increased from 44.2% (414/949) of those surveyed in 2009 to three-quarters (961/1278, 75.19%) in 2015. This finding is consistent with the increasing

prevalence rates of body image distress in countries such as the United States [1]. Furthermore, although there are some suggestions in the literature that concern regarding body image has increased in Australia [84-86], this is the first known study to report changes across these 3 time points using web-based samples. Our data indicate that the issue is much more prevalent. This increasing prevalence of body image distress corresponds with the rise of social media. During the time frame of the survey, Instagram was launched in 2010, Snapchat was released in 2011, and TikTok was released in 2016. As photographs and videos are central to the use of these platforms, and previous research has shown an association between body image distress and social media use, perhaps this increased level of distress has occurred in parallel with the rise of social media [14,15].

This rise in prevalence is particularly concerning, given our findings that body image distress was associated with increased levels of current dieting, poorer self-perception of weight, higher psychological distress, increased alcohol or other substance misuse, and poorer personal well-being. This is consistent with previous research where higher body image dissatisfaction directly correlated with poor mental health-related quality of life and psychosocial functioning [17]. Furthermore, literature examining body appreciation has reported associations with greater subjective happiness [56] and fewer days of feeling mentally or physically unhealthy [87]. It is unclear whether these factors are precipitating factors or consequences of body image distress. As 24% of the variability in body image distress was accounted for by these factors, future research could endeavor to explore what other factors are potentially missing from this model that also contribute to distress, such as social media use or history of disordered eating. Overall, when considering the rise of body image distress and its link to poorer psychosocial outcomes, a sharper focus on this area is needed.

Another important finding was that in our 2015 international sample, more than half of the participants' BMI in the overweight or obese range (59.34%), with men reporting higher levels of obesity than women. This in itself is concerning, as obesity is considered one of the greatest health epidemics worldwide [88-91]. Furthermore, our findings demonstrated a notable sex difference concerning how men and women perceive their weight compared with their self-reported weight as measured by BMI. Specifically, despite a higher percentage of men having an overweight or obese BMI, more women (772/3339, 23.12%) considered themselves to be very overweight than men (351/2176, 16.13%). Although more women in this sample were in a healthy weight range (women: 1344/3264, 41.19%; men: 671/2115, 31.73%), only one-third of women believed they were *about the right weight*. This is consistent with data from previous studies demonstrating that women are more likely to perceive themselves as overweight compared with men [35,92-95]. In the literature, possible factors that may contribute to this discrepancy include self-esteem [96], sociocultural influences, and expectations [97-99].

Interventions in this area are relatively unexplored, particularly those targeting both men and women. Evidence-based interventions include self-monitoring, cognitive restructuring, exposure exercises, fitness training, mirror work [100], media literacy, self-esteem enhancement, and psychoeducation.

However, these interventions only achieve minor improvements in body image [101,102]. Interventions with a greater focus on stress management training, cognitions, and negative body image causes appear to be more effective [101]. There is some evidence that self-compassion training can be beneficial for weight loss, nutrition behaviors, eating behaviors, and body image [103]. This training focuses on promoting self-worth, creating a more positive body image, and decreasing body dissatisfaction, and may be the way forward to improve health outcomes in distressed individuals.

These discrepancies between actual and perceived weight were not only a function of sex but also of age. Our 2015 survey results indicated that 57.77% of young people were in the healthy BMI weight range—the highest endorsement across all age groups. Despite this, three-quarters of young people reported body image distress. Again, this was the highest endorsement across all age groups. As participants aged, their BMI increased—with far more in the overweight and obese categories. However, the rate of body image distress declined as participants aged, as did the influence of weight or shape on how they viewed themselves as a person. Previous literature supports this phenomenon [104-107]. This change could be attributed to a shift in body comparisons with age-appropriate peers [8], less cultural fat phobia [105-107], or a focus on body function rather than body appearance [108].

Furthermore, research has theorized that people's preoccupation and desire to change their body weight via dieting behavior becomes less salient with age [26]. Interestingly, our survey results demonstrated that approximately three-quarters of people reported they were not dieting, and this proportion remained relatively stable across each age group. Thus, although body image distress decreases with age, in line with the Webster and Tiggemann study [26], our contradictory finding is that dieting behaviors remain relatively consistent. Further research is needed to examine whether this is explained by the changes in people's reasons for dieting as they age. For example, older people might be dieting for health or medical reasons rather than because of their body image concerns.

Implications for Policy and Practice

This study supports the clear link between body image distress and poorer health and well-being [28]. Prevalence rates of body image dissatisfaction have increased worldwide in the past 30 years [109]. Our research shows that this prevalence is 3 in 4 young people when using a web-based survey methodology. These are compelling statistics. In 2019, the Australian government invested US \$146 million into the prevention, detection, assessment, and treatment of eating disorders [110]. Although this is timely, our findings on the inverse relationship between individual distress and psychosocial outcomes make a strong case for the need for prevention and early intervention before eating disorders emerge. This may include more comprehensive assessment when accessing health services and the use of health information technologies to improve support services [111]. Given our findings, such interventions may benefit from targeting across sex and age.

Previously, body image distress was thought to result from the discrepancy between actual and perceived body image.

However, our results suggest that the rates of body image distress within some age groups far exceed the proportion of the population who experience a discrepancy, indicative of other factors contributing to distress. For example, research is needed to determine whether these results are related to the considerable increases in the use and availability of social media in the past decade [14]. Exposure to social media, particularly in an individual's formative years, could have a considerable impact on a young person's sense of self, quality of life, and body image than is currently known. Some studies have demonstrated that the use of highly visual social media such as Instagram or Snapchat is linked to upward social comparison and internalization of symptoms of body dysmorphia, resulting in increased body image distress [112-115]. The more time spent on social media, the more significant the body image concerns [116-118]. Photograph-based activity (eg, posting photographs and viewing or making comments on others' photographs) is particularly salient in contributing to body dissatisfaction and disordered eating [14]. As the data in this study were from a web-based sample recruited using social media channels, the results may reflect the experience of body image distress of a web-based population, who may be more likely to be using other platforms such as Instagram.

Body image distress and dieting behaviors are well understood to be risk factors for disordered eating and the development of an eating disorder. Our results indicate a relationship between dieting behavior, psychological distress, and the self-perception of weight, in addition to alcohol and/or other substance use and well-being impact on body image distress. The triangulation of items such as those used in this survey (BMI vs distress vs dieting status, or BMI vs distress vs perceptions of weight) may be beneficial as a brief screening tool. Given the burden of completing lengthy psychometrics, how these brief screeners compare with lengthier eating disorder questionnaires should be explored.

Strengths and Limitations

A key strength of this study is that it is one of the largest samples to date, providing data on weight, perceptions of weight, dieting status, body image distress, and health and well-being. As outlined in the main report's executive summary [69], a key limitation of the research is the non-epidemiological nature of the web-based research; targeting efforts were made for age, sex, and by country to address this. Although most individuals in the participating countries have widespread internet access [119], this study will also yield some level of internet bias, in that those who do not frequently access the internet or social media recruitment websites may not have participated. However, a major advantage to using a web-based surveying methodology

is that previous research has found that it is associated with increased disclosure of sensitive information [120], such as the items asked in this survey. There is also the possibility of avidity bias occurring, as those with a greater interest in the subject may be more likely to participate [121]. However, overall, this research remains highly relevant, as it is the interactions between variables, not merely the statistical frequencies, that generate meaningful information. Furthermore, as we move further into the 21st century, web-based questionnaires may become more the norm than the exception.

Where possible, measures that have been tested for reliability and validity across general populations worldwide were used. For example, the K10 is the standard tool used to measure distress in Australia's National Survey of Mental Health and Wellbeing (Burgess et al [122]) and is used widely in international studies. The BMI [121] is the most recommended and widely used tool for classifying weight range in adults [123]. However, the use of BMI is limited and has several deficiencies as a measure of obesity [124]. BMI is not a reliable reflection of health status, does not accurately reflect changes that occur with age, cannot account for muscle mass, and is a poor indicator of body fat percentage [123].

Furthermore, the participants were asked to self-report their height and weight. Responses may have been subject to bias as BMI is often calculated with overestimated height and underestimated weight data [125]. Furthermore, owing to the breadth and sheer size of the Global Health and Wellbeing Survey 2015, validated measures for eating disorders such as Eating Disorder Inventory, third edition [126] and the Eating Attitudes Test [127] were not viable to use. Instead, we asked brief questions (all adapted from established literature) on areas such as dieting status, perceptions of weight, and body image distress. Overall, the results are still meaningful. Although they may not be fully representative of the populations with eating disorders, they demonstrate a clear link between body image distress and health and well-being concerns.

Conclusions

This research demonstrates the significant interplay between body image distress and psychosocial risk factors, including currently dieting, worse perceptions of weight, elevated psychological distress, increased alcohol or other substance misuse, and poorer personal well-being. Considering that an increasing number of young people are experiencing body image distress, body image should be closely monitored, given its association with poorer health outcomes. Further research into tailored intervention and prevention strategies for those experiencing any level of body image distress, obesity, eating disorders, and other health-related concerns is needed.

Acknowledgments

The Global Health and Wellbeing 2015 Survey was commissioned by the Movember Foundation and conducted by the University of Sydney's Brain and Mind Centre and the Young and Well Cooperative Research Center (Young and Well Cooperative Research Centre: 2011-2016). The authors would like to acknowledge the respondents who consented to participate on the web in the Global Health and Wellbeing 2015 Survey; the international consortia for the Global Health and Wellbeing 2015 Survey, including Professor Sagar Parikh, Professor Richard Porter, Professor Jan Scott, and Dr Michael Rovito; the Movember Foundation

(Australia) lead on the project Therese Fitzpatrick; and the Brain and Mind Centre team: Victoria Baldwin, Lisa Whittle, Django White, Laura Ospina Pinillos, Sarah Piper, Hannah Yee, and Frank Iorfino. The authors would also like to acknowledge the Young and Well CRC and *headspace* for the 2012 and 2009 data sets. The Movember Foundation commissioned this research. This project is also supported by philanthropic funding, for which donors are families affected by mental illness who wish to remain anonymous. This study was also funded by a National Health and Medical Research Council Australia Fellowship (No. 511921, awarded to IH).

Conflicts of Interest

IH was an inaugural Commissioner on Australia's National Mental Health Commission (2012-18). He is the codirector of Health and Policy at the Brain and Mind Centre (University of Sydney, Australia). The Brain and Mind Centre operates early-intervention youth services at Camperdown under contract to *headspace*. IH has previously led community-based and pharmaceutical industry-supported (Wyeth, Eli Lilly, Servier, Pfizer, and AstraZeneca) projects focused on the identification and better management of anxiety and depression. He was a member of the Medical Advisory Panel for Medibank Private until October 2017, a board member of Psychosis Australia Trust, and a member of the Veterans Mental Health Clinical Reference group. He is the chief scientific advisor to and a 5% equity shareholder in InnoWell Pty Ltd. InnoWell was formed by the University of Sydney (45% equity) and PwC (Australia; 45% equity) to deliver the Aus \$30 (US \$22.2) million Australian government-funded Project Synergy (2017-20; a 3-year program for the transformation of mental health services) and lead transformation of mental health services internationally through the use of innovative technologies. JB is chair of the National Advisory Council for Open Arms, Veterans, and Families Counseling Service. She is a well-being and digital health consultant to Bupa, a member of the Veterans Mental Health Clinical Reference group, and a chief investigator and author of the Defense and Veterans Transition and Well-being Study. She is the founder of and an equity shareholder in InnoWell. She is a professor of Social Innovation and Chair of the Center for Mental Health at Swinburne University and an adjunct professor of Social Impact and Entrepreneurship at Royal Melbourne Institute of Technology. TD is now the director of Research & Insights at the Australian Digital Health Agency. The other authors have no conflicts of interest to disclose.

Multimedia Appendix 1

Frequency statistics, chi-square, and analysis of variance comparing body image items by age and sex.

[\[DOCX File , 24 KB - formative_v5i11e25329_app1.docx \]](#)

Multimedia Appendix 2

Linear regression of body image distress by age group.

[\[DOCX File , 17 KB - formative_v5i11e25329_app2.docx \]](#)

Multimedia Appendix 3

Pearson correlations of body image items with health and well-being measures for each age group.

[\[DOCX File , 15 KB - formative_v5i11e25329_app3.docx \]](#)

References

1. Adams G, Turner H, Bucks R. The experience of body dissatisfaction in men. *Body Image* 2005 Sep;2(3):271-283. [doi: [10.1016/j.bodyim.2005.05.004](https://doi.org/10.1016/j.bodyim.2005.05.004)] [Medline: [18089194](https://pubmed.ncbi.nlm.nih.gov/18089194/)]
2. Tiggemann M, Pickering AS. Role of television in adolescent women's body dissatisfaction and drive for thinness. *Int J Eat Disord* 1996 Sep;20(2):199-203. [doi: [10.1002/\(SICD\)1098-108X\(199609\)20:2<199::AID-EAT11>3.0.CO;2-Z](https://doi.org/10.1002/(SICD)1098-108X(199609)20:2<199::AID-EAT11>3.0.CO;2-Z)] [Medline: [8863073](https://pubmed.ncbi.nlm.nih.gov/8863073/)]
3. Cash TF, Henry PE. Women's body images: the results of a national survey in the U.S.A. *Sex Roles* 1995 Jul;33(1-2):19-28. [doi: [10.1007/bf01547933](https://doi.org/10.1007/bf01547933)]
4. Peat CM, Peyerl NL, Muehlenkamp JJ. Body image and eating disorders in older adults: a review. *J Gen Psychol* 2008 Oct;135(4):343-358. [doi: [10.3200/GENP.135.4.343-358](https://doi.org/10.3200/GENP.135.4.343-358)] [Medline: [18959226](https://pubmed.ncbi.nlm.nih.gov/18959226/)]
5. McCabe MP, Ricciardelli LA. Body image dissatisfaction among males across the lifespan: a review of past literature. *J Psychosom Res* 2004 Jun;56(6):675-685. [doi: [10.1016/S0022-3999\(03\)00129-6](https://doi.org/10.1016/S0022-3999(03)00129-6)] [Medline: [15193964](https://pubmed.ncbi.nlm.nih.gov/15193964/)]
6. Cash TF. Body image: past, present, and future. *Body Image* 2004 Jan;1(1):1-5. [doi: [10.1016/S1740-1445\(03\)00011-1](https://doi.org/10.1016/S1740-1445(03)00011-1)] [Medline: [18089136](https://pubmed.ncbi.nlm.nih.gov/18089136/)]
7. Cash TF, Szymanski ML. The development and validation of the Body-Image Ideals Questionnaire. *J Pers Assess* 1995 Jun;64(3):466-477. [doi: [10.1207/s15327752jpa6403_6](https://doi.org/10.1207/s15327752jpa6403_6)] [Medline: [16367722](https://pubmed.ncbi.nlm.nih.gov/16367722/)]
8. Grogan S. *Body Image: Understanding Body Dissatisfaction in Men Women and Children (Third Edition)*. Oxfordshire England: Routledge; Sep 29, 2016.
9. Mitchison D, Hay P, Griffiths S, Murray SB, Bentley C, Gratwick-Sarll K, et al. Disentangling body image: the relative associations of overvaluation, dissatisfaction, and preoccupation with psychological distress and eating disorder behaviors

- in male and female adolescents. *Int J Eat Disord* 2017 Feb;50(2):118-126 [FREE Full text] [doi: [10.1002/eat.22592](https://doi.org/10.1002/eat.22592)] [Medline: [27539911](https://pubmed.ncbi.nlm.nih.gov/27539911/)]
10. Glashouwer KA, Bennik EC, de Jong PJ, Spruyt A. Implicit measures of actual versus ideal body image: relations with self-reported body dissatisfaction and dieting behaviors. *Cognit Ther Res* 2018;42(5):622-635 [FREE Full text] [doi: [10.1007/s10608-018-9917-6](https://doi.org/10.1007/s10608-018-9917-6)] [Medline: [30237650](https://pubmed.ncbi.nlm.nih.gov/30237650/)]
 11. Kelley CC, Neufeld JM, Musher-Eizenman DR. Drive for thinness and drive for muscularity: opposite ends of the continuum or separate constructs? *Body Image* 2010 Jan;7(1):74-77. [doi: [10.1016/j.bodyim.2009.09.008](https://doi.org/10.1016/j.bodyim.2009.09.008)] [Medline: [19944659](https://pubmed.ncbi.nlm.nih.gov/19944659/)]
 12. 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/)]
 13. Tiller E, Fildes J, Hall S. Youth survey report 2020. Mission Australia. URL: <https://apo.org.au/node/309542> [accessed 2021-08-27]
 14. Holland G, Tiggemann M. A systematic review of the impact of the use of social networking sites on body image and disordered eating outcomes. *Body Image* 2016 Jun;17:100-110. [doi: [10.1016/j.bodyim.2016.02.008](https://doi.org/10.1016/j.bodyim.2016.02.008)] [Medline: [26995158](https://pubmed.ncbi.nlm.nih.gov/26995158/)]
 15. Groesz LM, Levine MP, Murnen SK. The effect of experimental presentation of thin media images on body satisfaction: a meta-analytic review. *Int J Eat Disord* 2002 Jan;31(1):1-16. [doi: [10.1002/eat.10005](https://doi.org/10.1002/eat.10005)] [Medline: [11835293](https://pubmed.ncbi.nlm.nih.gov/11835293/)]
 16. Weinberger N, Kersting A, Riedel-Heller SG, Luck-Sikorski C. Body dissatisfaction in individuals with obesity compared to normal-weight individuals: a systematic review and meta-analysis. *Obes Facts* 2016;9(6):424-441 [FREE Full text] [doi: [10.1159/000454837](https://doi.org/10.1159/000454837)] [Medline: [28013298](https://pubmed.ncbi.nlm.nih.gov/28013298/)]
 17. Wilson RE, Latner JD, Hayashi K. More than just body weight: the role of body image in psychological and physical functioning. *Body Image* 2013 Sep;10(4):644-647. [doi: [10.1016/j.bodyim.2013.04.007](https://doi.org/10.1016/j.bodyim.2013.04.007)] [Medline: [23726517](https://pubmed.ncbi.nlm.nih.gov/23726517/)]
 18. Friedman K, Reichmann S, Costanzo P, Musante G. Body image partially mediates the relationship between obesity and psychological distress. *Obes Res* 2002 Jan;10(1):33-41 [FREE Full text] [doi: [10.1038/oby.2002.5](https://doi.org/10.1038/oby.2002.5)] [Medline: [11786599](https://pubmed.ncbi.nlm.nih.gov/11786599/)]
 19. Kantanista A, Król-Zielińska M, Borowiec J, Osirski W. Is underweight associated with more positive body image? Results of a cross-sectional study in adolescent girls and boys. *Span J Psychol* 2017 Feb 09;20:E8. [doi: [10.1017/sjp.2017.4](https://doi.org/10.1017/sjp.2017.4)] [Medline: [28181473](https://pubmed.ncbi.nlm.nih.gov/28181473/)]
 20. Sarwer DB, Thompson JK, Cash TF. Body image and obesity in adulthood. *Psychiatr Clin North Am* 2005 Mar;28(1):69-87, viii. [doi: [10.1016/j.psc.2004.09.002](https://doi.org/10.1016/j.psc.2004.09.002)] [Medline: [15733612](https://pubmed.ncbi.nlm.nih.gov/15733612/)]
 21. Allaz A, Bernstein M, Rouget P, Archinard M, Morabia A. Body weight preoccupation in middle-age and ageing women: a general population survey. *Int J Eat Disord* 1998 Apr;23(3):287-294. [doi: [10.1002/\(sici\)1098-108x\(199804\)23:3<287::aid-eat6>3.0.co;2-f](https://doi.org/10.1002/(sici)1098-108x(199804)23:3<287::aid-eat6>3.0.co;2-f)] [Medline: [9547663](https://pubmed.ncbi.nlm.nih.gov/9547663/)]
 22. Thompson JK, Stice E. Thin-ideal internalization: mounting evidence for a new risk factor for body-image disturbance and eating pathology. *Curr Dir Psychol Sci* 2001 Oct 1;10(5):181-183. [doi: [10.1111/1467-8721.00144](https://doi.org/10.1111/1467-8721.00144)]
 23. Watkins JA, Christie C, Chally P. Relationship between body image and body mass index in college men. *J Am Coll Health* 2008;57(1):95-100. [doi: [10.3200/JACH.57.1.95-100](https://doi.org/10.3200/JACH.57.1.95-100)] [Medline: [18682351](https://pubmed.ncbi.nlm.nih.gov/18682351/)]
 24. Anderson D, Lundgren J, Shapiro J, Paulosky C. Weight goals in a college-age population. *Obes Res* 2003 Feb;11(2):274-278 [FREE Full text] [doi: [10.1038/oby.2003.41](https://doi.org/10.1038/oby.2003.41)] [Medline: [12582224](https://pubmed.ncbi.nlm.nih.gov/12582224/)]
 25. Runfola CD, Von Holle A, Peat CM, Gagne DA, Brownley KA, Hofmeier SM, et al. Characteristics of women with body size satisfaction at midlife: results of the Gender and Body Image (GABI) Study. *J Women Aging* 2013;25(4):287-304 [FREE Full text] [doi: [10.1080/08952841.2013.816215](https://doi.org/10.1080/08952841.2013.816215)] [Medline: [24116991](https://pubmed.ncbi.nlm.nih.gov/24116991/)]
 26. Webster J, Tiggemann M. The relationship between women's body satisfaction and self-image across the life span: the role of cognitive control. *J Genet Psychol* 2003 Jun;164(2):241-252. [doi: [10.1080/00221320309597980](https://doi.org/10.1080/00221320309597980)] [Medline: [12856818](https://pubmed.ncbi.nlm.nih.gov/12856818/)]
 27. Rand CS, Wright BA. Continuity and change in the evaluation of ideal and acceptable body sizes across a wide age span. *Int J Eat Disord* 2000 Jul;28(1):90-100. [doi: [10.1002/\(sici\)1098-108x\(200007\)28:1<90::aid-eat11>3.0.co;2-p](https://doi.org/10.1002/(sici)1098-108x(200007)28:1<90::aid-eat11>3.0.co;2-p)] [Medline: [10800018](https://pubmed.ncbi.nlm.nih.gov/10800018/)]
 28. van den Berg PA, Mond J, Eisenberg M, Ackard D, Neumark-Sztainer D. The link between body dissatisfaction and self-esteem in adolescents: similarities across gender, age, weight status, race/ethnicity, and socioeconomic status. *J Adolesc Health* 2010 Sep;47(3):290-296 [FREE Full text] [doi: [10.1016/j.jadohealth.2010.02.004](https://doi.org/10.1016/j.jadohealth.2010.02.004)] [Medline: [20708569](https://pubmed.ncbi.nlm.nih.gov/20708569/)]
 29. Cruz-Sáez S, Pascual A, Salaberria K, Echeburúa E. Normal-weight and overweight female adolescents with and without extreme weight-control behaviours: emotional distress and body image concerns. *J Health Psychol* 2015 Jun;20(6):730-740. [doi: [10.1177/1359105315580214](https://doi.org/10.1177/1359105315580214)] [Medline: [26032790](https://pubmed.ncbi.nlm.nih.gov/26032790/)]
 30. Dooley B, Fitzgerald A, Giollabhui NM. The risk and protective factors associated with depression and anxiety in a national sample of Irish adolescents. *Ir J Psychol Med* 2015 Mar;32(1):93-105. [doi: [10.1017/ipm.2014.83](https://doi.org/10.1017/ipm.2014.83)] [Medline: [30185277](https://pubmed.ncbi.nlm.nih.gov/30185277/)]
 31. Ferguson W, Goode M, Walsh A, Gavin P, Butler K. Evaluation of 4 weeks' neonatal antiretroviral prophylaxis as a component of a prevention of mother-to-child transmission program in a resource-rich setting. *Pediatr Infect Dis J* 2011 May;30(5):408-412. [doi: [10.1097/INF.0b013e31820614bd](https://doi.org/10.1097/INF.0b013e31820614bd)] [Medline: [21266939](https://pubmed.ncbi.nlm.nih.gov/21266939/)]
 32. Kostanski M, Gullone E. Adolescent body image dissatisfaction: relationships with self-esteem, anxiety, and depression controlling for body mass. *J Child Psychol Psychiatry* 1998 Feb;39(2):255-262. [Medline: [9669238](https://pubmed.ncbi.nlm.nih.gov/9669238/)]

33. Koronczi B, Kökönyei G, Urbán R, Kun B, Pápay O, Nagygyörgy K, et al. The mediating effect of self-esteem, depression and anxiety between satisfaction with body appearance and problematic internet use. *Am J Drug Alcohol Abuse* 2013 Jul;39(4):259-265. [doi: [10.3109/00952990.2013.803111](https://doi.org/10.3109/00952990.2013.803111)] [Medline: [23841866](https://pubmed.ncbi.nlm.nih.gov/23841866/)]
34. Bélanger M, Marcotte D. Étude longitudinale du lien entre les changements vécus durant la transition primaire-secondaire et les symptômes dépressifs des adolescents. *Can J Behav Sci/Revue canadienne des sciences du comportement* 2013;45(2):159-172. [doi: [10.1037/a0025010](https://doi.org/10.1037/a0025010)]
35. Wardle J, Johnson F. Weight and dieting: examining levels of weight concern in British adults. *Int J Obes Relat Metab Disord* 2002 Aug;26(8):1144-1149. [doi: [10.1038/sj.ijo.0802046](https://doi.org/10.1038/sj.ijo.0802046)] [Medline: [12119582](https://pubmed.ncbi.nlm.nih.gov/12119582/)]
36. Stice E, Bearman SK. Body-image and eating disturbances prospectively predict increases in depressive symptoms in adolescent girls: a growth curve analysis. *Dev Psychol* 2001 Sep;37(5):597-607. [doi: [10.1037//0012-1649.37.5.597](https://doi.org/10.1037//0012-1649.37.5.597)] [Medline: [11552756](https://pubmed.ncbi.nlm.nih.gov/11552756/)]
37. Wichstrøm L, von Soest T. Reciprocal relations between body satisfaction and self-esteem: a large 13-year prospective study of adolescents. *J Adolesc* 2016 Feb;47:16-27. [doi: [10.1016/j.adolescence.2015.12.003](https://doi.org/10.1016/j.adolescence.2015.12.003)] [Medline: [26713411](https://pubmed.ncbi.nlm.nih.gov/26713411/)]
38. Brausch AM, Muehlenkamp JJ. Body image and suicidal ideation in adolescents. *Body Image* 2007 Jun;4(2):207-212. [doi: [10.1016/j.bodyim.2007.02.001](https://doi.org/10.1016/j.bodyim.2007.02.001)] [Medline: [18089266](https://pubmed.ncbi.nlm.nih.gov/18089266/)]
39. Brausch AM, Gutierrez PM. The role of body image and disordered eating as risk factors for depression and suicidal ideation in adolescents. *Suicide Life Threat Behav* 2009 Feb;39(1):58-71. [doi: [10.1521/suli.2009.39.1.58](https://doi.org/10.1521/suli.2009.39.1.58)] [Medline: [19298151](https://pubmed.ncbi.nlm.nih.gov/19298151/)]
40. Orbach I, Stein D, Shan-Sela M, Har-Even D. Body attitudes and body experiences in suicidal adolescents. *Suicide Life Threat Behav* 2001;31(3):237-249. [doi: [10.1521/suli.31.3.237.24250](https://doi.org/10.1521/suli.31.3.237.24250)] [Medline: [11577910](https://pubmed.ncbi.nlm.nih.gov/11577910/)]
41. Orbach I, Gilboa-Schechtman E, Sheffer A, Meged S, Har-Even D, Stein D. Negative bodily self in suicide attempters. *Suicide Life Threat Behav* 2006 Apr;36(2):136-153. [doi: [10.1521/suli.2006.36.2.136](https://doi.org/10.1521/suli.2006.36.2.136)] [Medline: [16704320](https://pubmed.ncbi.nlm.nih.gov/16704320/)]
42. Miotto P, De Coppi M, Frezza M, Preti A. Eating disorders and suicide risk factors in adolescents: an Italian community-based study. *J Nerv Ment Dis* 2003 Jul;191(7):437-443. [doi: [10.1097/01.NMD.0000081590.91326.8B](https://doi.org/10.1097/01.NMD.0000081590.91326.8B)] [Medline: [12891090](https://pubmed.ncbi.nlm.nih.gov/12891090/)]
43. Forman M, Davis WN. Characteristics of middle-aged women in inpatient treatment for eating disorders. *Eat Disord* 2005;13(3):231-243. [doi: [10.1080/10640260590932841](https://doi.org/10.1080/10640260590932841)] [Medline: [16864530](https://pubmed.ncbi.nlm.nih.gov/16864530/)]
44. Klemchuk HP, Hutchinson CB, Frank RI. Body dissatisfaction and eating-related problems on the college campus: usefulness of the Eating Disorder Inventory with a nonclinical population. *J Counsel Psychol* 1990 Jul;37(3):297-305. [doi: [10.1037/0022-0167.37.3.297](https://doi.org/10.1037/0022-0167.37.3.297)]
45. Tylka T. The relation between body dissatisfaction and eating disorder symptomatology: an analysis of moderating variables. *J Counsel Psychol* 2004 Apr;51(2):178-191. [doi: [10.1037/0022-0167.51.2.178](https://doi.org/10.1037/0022-0167.51.2.178)]
46. Griffiths S, Hay P, Mitchison D, Mond JM, McLean SA, Rodgers B, et al. Sex differences in the relationships between body dissatisfaction, quality of life and psychological distress. *Aust N Z J Public Health* 2016 Dec;40(6):518-522. [doi: [10.1111/1753-6405.12538](https://doi.org/10.1111/1753-6405.12538)] [Medline: [27372301](https://pubmed.ncbi.nlm.nih.gov/27372301/)]
47. Kendzor DE, Adams CE, Stewart DW, Baillie LE, Copeland AL. Cigarette smoking is associated with body shape concerns and bulimia symptoms among young adult females. *Eat Behav* 2009 Jan;10(1):56-58 [FREE Full text] [doi: [10.1016/j.eatbeh.2008.10.012](https://doi.org/10.1016/j.eatbeh.2008.10.012)] [Medline: [19171320](https://pubmed.ncbi.nlm.nih.gov/19171320/)]
48. Nelson M, Lust K, Story M, Ehlinger E. Alcohol use, eating patterns, and weight behaviors in a university population. *Am J Health Behav* 2009;33(3):227-237. [doi: [10.5993/ajhb.33.3.1](https://doi.org/10.5993/ajhb.33.3.1)] [Medline: [19063644](https://pubmed.ncbi.nlm.nih.gov/19063644/)]
49. Corning AF, Krumm AJ, Smitham LA. Differential social comparison processes in women with and without eating disorder symptoms. *J Counsel Psychol* 2006;53(3):338-349. [doi: [10.1037/0022-0167.53.3.338](https://doi.org/10.1037/0022-0167.53.3.338)]
50. Green SP, Pritchard ME. Predictors of body image dissatisfaction in adult men and women. *Soc Behav Pers* 2003 Jan 01;31(3):215-222. [doi: [10.2224/sbp.2003.31.3.215](https://doi.org/10.2224/sbp.2003.31.3.215)]
51. Alleva JM, Martijn C, Veldhuis J, Tylka TL. A Dutch translation and validation of the Body Appreciation Scale-2: an investigation with female university students in the Netherlands. *Body Image* 2016 Dec;19:44-48. [doi: [10.1016/j.bodyim.2016.08.008](https://doi.org/10.1016/j.bodyim.2016.08.008)] [Medline: [27598764](https://pubmed.ncbi.nlm.nih.gov/27598764/)]
52. Dalley SE, Vidal J. Optimism and positive body image in women: the mediating role of the feared fat self. *Pers Individ Dif* 2013 Sep;55(5):465-468. [doi: [10.1016/j.paid.2013.04.006](https://doi.org/10.1016/j.paid.2013.04.006)]
53. Marta-Simões J, Ferreira C, Mendes AL. Exploring the effect of external shame on body appreciation among Portuguese young adults: the role of self-compassion. *Eat Behav* 2016 Dec;23:174-179. [doi: [10.1016/j.eatbeh.2016.10.006](https://doi.org/10.1016/j.eatbeh.2016.10.006)] [Medline: [27816855](https://pubmed.ncbi.nlm.nih.gov/27816855/)]
54. Raque-Bogdan TL, Piontkowski S, Hui K, Ziemer KS, Garriott PO. Self-compassion as a mediator between attachment anxiety and body appreciation: an exploratory model. *Body Image* 2016 Dec;19:28-36. [doi: [10.1016/j.bodyim.2016.08.001](https://doi.org/10.1016/j.bodyim.2016.08.001)] [Medline: [27597725](https://pubmed.ncbi.nlm.nih.gov/27597725/)]
55. Stokes R, Frederick-Recascino C. Women's perceived body image: relations with personal happiness. *J Women Aging* 2003;15(1):17-29. [doi: [10.1300/J074v15n01_03](https://doi.org/10.1300/J074v15n01_03)] [Medline: [12678183](https://pubmed.ncbi.nlm.nih.gov/12678183/)]
56. Swami V, Ng S. Factor structure and psychometric properties of the Body Appreciation Scale-2 in university students in Hong Kong. *Body Image* 2015 Sep;15:68-71. [doi: [10.1016/j.bodyim.2015.06.004](https://doi.org/10.1016/j.bodyim.2015.06.004)] [Medline: [26189168](https://pubmed.ncbi.nlm.nih.gov/26189168/)]
57. Swami V, Ng S, Barron D. Translation and psychometric evaluation of a Standard Chinese version of the Body Appreciation Scale-2. *Body Image* 2016 Sep;18:23-26. [doi: [10.1016/j.bodyim.2016.04.005](https://doi.org/10.1016/j.bodyim.2016.04.005)] [Medline: [27236474](https://pubmed.ncbi.nlm.nih.gov/27236474/)]

58. Swami V, Tran US, Stieger S, Voracek M, The YouBeauty.com Team 2015. Associations between women's body image and happiness: results of the YouBeauty.com Body Image Survey (YBIS). *J Happiness Stud* 2014 Apr 26;16(3):705-718. [doi: [10.1007/s10902-014-9530-7](https://doi.org/10.1007/s10902-014-9530-7)]
59. Swami V, Weis L, Barron D, Furnham A. Positive body image is positively associated with hedonic (emotional) and eudaimonic (psychological and social) well-being in British adults. *J Soc Psychol* 2018;158(5):541-552. [doi: [10.1080/00224545.2017.1392278](https://doi.org/10.1080/00224545.2017.1392278)] [Medline: [29053404](https://pubmed.ncbi.nlm.nih.gov/29053404/)]
60. Tylka TL, Van Diest AM. The Intuitive Eating Scale-2: item refinement and psychometric evaluation with college women and men. *J Couns Psychol* 2013 Jan;60(1):137-153. [doi: [10.1037/a0030893](https://doi.org/10.1037/a0030893)] [Medline: [23356469](https://pubmed.ncbi.nlm.nih.gov/23356469/)]
61. Wasylikiw L, MacKinnon AL, MacLellan AM. Exploring the link between self-compassion and body image in university women. *Body Image* 2012 Mar;9(2):236-245. [doi: [10.1016/j.bodyim.2012.01.007](https://doi.org/10.1016/j.bodyim.2012.01.007)] [Medline: [22406200](https://pubmed.ncbi.nlm.nih.gov/22406200/)]
62. Barker ET, Galambos NL. Body dissatisfaction of adolescent girls and boys:: risk and resource factors. *J Early Adolesc* 2003 May 1;23(2):141-165. [doi: [10.1177/0272431603023002002](https://doi.org/10.1177/0272431603023002002)]
63. Bearman SK, Martinez E, Stice E, Presnell K. The skinny on body dissatisfaction: a longitudinal study of adolescent girls and boys. *J Youth Adolesc* 2006 Apr;35(2):217-229 [FREE Full text] [doi: [10.1007/s10964-005-9010-9](https://doi.org/10.1007/s10964-005-9010-9)] [Medline: [16912810](https://pubmed.ncbi.nlm.nih.gov/16912810/)]
64. Stice E, Whitenton K. Risk factors for body dissatisfaction in adolescent girls: a longitudinal investigation. *Dev Psychol* 2002 Sep;38(5):669-678. [doi: [10.1037//0012-1649.38.5.669](https://doi.org/10.1037//0012-1649.38.5.669)] [Medline: [12220046](https://pubmed.ncbi.nlm.nih.gov/12220046/)]
65. Close S, Smaldone A, Fennoy I, Reame N, Grey M. Using information technology and social networking for recruitment of research participants: experience from an exploratory study of pediatric Klinefelter syndrome. *J Med Internet Res* 2013 Mar 19;15(3):e48 [FREE Full text] [doi: [10.2196/jmir.2286](https://doi.org/10.2196/jmir.2286)] [Medline: [23512442](https://pubmed.ncbi.nlm.nih.gov/23512442/)]
66. Merolli M, Gray K, Martin-Sanchez F. Therapeutic affordances of social media: emergent themes from a global online survey of people with chronic pain. *J Med Internet Res* 2014 Dec 22;16(12):e284 [FREE Full text] [doi: [10.2196/jmir.3494](https://doi.org/10.2196/jmir.3494)] [Medline: [25533453](https://pubmed.ncbi.nlm.nih.gov/25533453/)]
67. Bhutta CB. Not by the book: Facebook as a sampling frame. *Sociol Methods Res* 2012 Mar 21;41(1):57-88. [doi: [10.1177/0049124112440795](https://doi.org/10.1177/0049124112440795)]
68. Biernacki P, Waldorf D. Snowball sampling: problems and techniques of chain referral sampling. *Sociol Methods Res* 1981 Nov;10(2):141-163. [doi: [10.1177/004912418101000205](https://doi.org/10.1177/004912418101000205)]
69. Burns J, Davenport T, Milton AI. Men's health and wellbeing in Australia, Canada, New Zealand, the United Kingdom and United States: findings from an international online pilot survey. The Movember Foundation. URL: https://cdn.movember.com/uploads/images/2017/Our%20Work/Publications/Mens_Health_Survey_Executive_Summary_FA.pdf [accessed 2021-08-27]
70. Blashill AJ, Wilhelm S. Body image distortions, weight, and depression in adolescent boys: longitudinal trajectories into adulthood. *Psychol Men Masc* 2014;15(4):445-451 [FREE Full text] [doi: [10.1037/a0034618](https://doi.org/10.1037/a0034618)] [Medline: [25383047](https://pubmed.ncbi.nlm.nih.gov/25383047/)]
71. Burns J, Christensen H, Luscombe G. Game on: exploring the impact of technologies on young men's mental health and wellbeing. Young and Well Cooperative Research Centre. URL: https://www.researchgate.net/publication/303812662_Game_On_Exploring_the_Impact_of_Technologies_on_Young_Men's_Mental_Health_and_Wellbeing [accessed 2021-08-27]
72. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003 Aug;35(8):1381-1395. [doi: [10.1249/01.MSS.0000078924.61453.FB](https://doi.org/10.1249/01.MSS.0000078924.61453.FB)] [Medline: [12900694](https://pubmed.ncbi.nlm.nih.gov/12900694/)]
73. Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, et al. Prevalence and treatment of mental disorders, 1990 to 2003. *N Engl J Med* 2005 Jun 16;352(24):2515-2523 [FREE Full text] [doi: [10.1056/NEJMsa043266](https://doi.org/10.1056/NEJMsa043266)] [Medline: [15958807](https://pubmed.ncbi.nlm.nih.gov/15958807/)]
74. Psychological distress. Australian Bureau of Statistics. 2012. URL: <https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4364.0.55.001~2014-15~Main%20Features~Psychological%20distress~16> [accessed 2021-09-08]
75. Lindelow M, Hardy R, Rodgers B. Development of a scale to measure symptoms of anxiety and depression in the general UK population: the psychiatric symptom frequency scale. *J Epidemiol Community Health* 1997 Oct;51(5):549-557 [FREE Full text] [doi: [10.1136/jech.51.5.549](https://doi.org/10.1136/jech.51.5.549)] [Medline: [9425466](https://pubmed.ncbi.nlm.nih.gov/9425466/)]
76. Von Korff M, Ustun T, Ormel J, Kaplan I, Simon GE. Self-report disability in an international primary care study of psychological illness. *J Clin Epidemiol* 1996 Mar;49(3):297-303. [doi: [10.1016/0895-4356\(95\)00512-9](https://doi.org/10.1016/0895-4356(95)00512-9)] [Medline: [8676177](https://pubmed.ncbi.nlm.nih.gov/8676177/)]
77. Personal wellbeing index. Australian Centre on Quality of Life. URL: <http://www.acqol.com.au/index> [accessed 2021-08-27]
78. Hills P, Argyle M. The Oxford Happiness Questionnaire: a compact scale for the measurement of psychological well-being. *Pers Individ Dif* 2002 Nov;33(7):1073-1082. [doi: [10.1016/s0191-8869\(01\)00213-6](https://doi.org/10.1016/s0191-8869(01)00213-6)]
79. Sinclair VG, Wallston KA. The development and psychometric evaluation of the Brief Resilient Coping Scale. *Assessment* 2004 Mar;11(1):94-101. [doi: [10.1177/1073191103258144](https://doi.org/10.1177/1073191103258144)] [Medline: [14994958](https://pubmed.ncbi.nlm.nih.gov/14994958/)]
80. Schuster T, Kessler R, Aseltine R. Supportive interactions, negative interactions, and depressed mood. *Am J Community Psychol* 1990 Jun;18(3):423-438. [doi: [10.1007/BF00938116](https://doi.org/10.1007/BF00938116)] [Medline: [2264558](https://pubmed.ncbi.nlm.nih.gov/2264558/)]
81. Kaplan RM, Chambers DA, Glasgow RE. Big data and large sample size: a cautionary note on the potential for bias. *Clin Transl Sci* 2014 Aug;7(4):342-346 [FREE Full text] [doi: [10.1111/cts.12178](https://doi.org/10.1111/cts.12178)] [Medline: [25043853](https://pubmed.ncbi.nlm.nih.gov/25043853/)]

82. Milton AC, Davenport TA, Iorfino F, Flego A, Burns JM, Hickie IB. Suicidal thoughts and behaviors and their associations with transitional life events in men and women: findings from an international web-based sample. *JMIR Ment Health* 2020 Sep 11;7(9):e18383 [FREE Full text] [doi: [10.2196/18383](https://doi.org/10.2196/18383)] [Medline: [32915160](https://pubmed.ncbi.nlm.nih.gov/32915160/)]
83. Milton AC, La Monica H, Dowling M, Yee H, Davenport T, Braunstein K, et al. Gambling and the role of resilience in an international online sample of current and ex-serving military personnel as compared to the general population. *J Gambli Stud* 2020 Jun;36(2):477-498. [doi: [10.1007/s10899-019-09900-w](https://doi.org/10.1007/s10899-019-09900-w)] [Medline: [31620927](https://pubmed.ncbi.nlm.nih.gov/31620927/)]
84. Buckley H, Dalton B, Fildes J, Ivancic L, Matkovic L, Perrens B, et al. Youth Survey 2012. Mission Australia. URL: <https://www.missionaustralia.com.au/publications/youth-survey/73-mission-australia-youth-survey-2012/file> [accessed 2021-08-27]
85. Carlisle E, Fildes J, Hall S, Hicking V, Perrens B, Plummer J. Youth Survey Report 2018. Mission Australia. URL: <https://apo.org.au/node/206361> [accessed 2021-08-27]
86. Mission Australia's 2014 Youth Survey Report. Mission Australia. URL: <https://www.ncsehe.edu.au/publications/mission-australia-youth-survey-2014/> [accessed 2021-08-27]
87. Muennig P, Jia H, Lee R, Lubetkin E. I think therefore I am: perceived ideal weight as a determinant of health. *Am J Public Health* 2008 Mar;98(3):501-506. [doi: [10.2105/AJPH.2007.114769](https://doi.org/10.2105/AJPH.2007.114769)] [Medline: [18235062](https://pubmed.ncbi.nlm.nih.gov/18235062/)]
88. 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/)]
89. Friedrich M. Global obesity epidemic worsening. *J Am Med Assoc* 2017 Aug 15;318(7):603. [doi: [10.1001/jama.2017.10693](https://doi.org/10.1001/jama.2017.10693)] [Medline: [28810033](https://pubmed.ncbi.nlm.nih.gov/28810033/)]
90. Wadden T, Womble L, Stunkard A, Anderson D. Psychosocial consequences of obesity and weight loss. In: Wadden TA, Stunkard AJ, editors. *Handbook of Obesity Treatment*. New York: Guilford Press; 2004:1-624.
91. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser* 2000;894:i-xii, 1-253. [Medline: [11234459](https://pubmed.ncbi.nlm.nih.gov/11234459/)]
92. Bish CL, Blanck HM, Maynard LM, Serdula MK, Thompson NJ, Khan LK. Health-related quality of life and weight loss among overweight and obese U.S. adults, 2001 to 2002. *Obesity (Silver Spring)* 2006 Nov;14(11):2042-2053 [FREE Full text] [doi: [10.1038/oby.2006.239](https://doi.org/10.1038/oby.2006.239)] [Medline: [17135622](https://pubmed.ncbi.nlm.nih.gov/17135622/)]
93. Chang VW, Christakis NA. Self-perception of weight appropriateness in the United States. *Am J Prev Med* 2003 May;24(4):332-339. [doi: [10.1016/s0749-3797\(03\)00020-5](https://doi.org/10.1016/s0749-3797(03)00020-5)] [Medline: [12726871](https://pubmed.ncbi.nlm.nih.gov/12726871/)]
94. Paeratakul S, Lovejoy J, Ryan D, Bray G. The relation of gender, race and socioeconomic status to obesity and obesity comorbidities in a sample of US adults. *Int J Obes Relat Metab Disord* 2002 Sep;26(9):1205-1210. [doi: [10.1038/sj.ijo.0802026](https://doi.org/10.1038/sj.ijo.0802026)] [Medline: [12187397](https://pubmed.ncbi.nlm.nih.gov/12187397/)]
95. Schieman S, Pudrovska T, Eccles R. Perceptions of body weight among older adults: analyses of the intersection of gender, race, and socioeconomic status. *J Gerontol B Psychol Sci Soc Sci* 2007 Nov;62(6):415-423. [doi: [10.1093/geronb/62.6.s415](https://doi.org/10.1093/geronb/62.6.s415)] [Medline: [18079430](https://pubmed.ncbi.nlm.nih.gov/18079430/)]
96. O'Dea JA, Abraham S. Improving the body image, eating attitudes, and behaviors of young male and female adolescents: a new educational approach that focuses on self-esteem. *Int J Eat Disord* 2000 Jul;28(1):43-57. [doi: [10.1002/\(sici\)1098-108x\(200007\)28:1<43::aid-eat6>3.0.co;2-d](https://doi.org/10.1002/(sici)1098-108x(200007)28:1<43::aid-eat6>3.0.co;2-d)] [Medline: [10800013](https://pubmed.ncbi.nlm.nih.gov/10800013/)]
97. Knauss C, Paxton SJ, Alsaker FD. Relationships amongst body dissatisfaction, internalisation of the media body ideal and perceived pressure from media in adolescent girls and boys. *Body Image* 2007 Dec;4(4):353-360. [doi: [10.1016/j.bodyim.2007.06.007](https://doi.org/10.1016/j.bodyim.2007.06.007)] [Medline: [18089281](https://pubmed.ncbi.nlm.nih.gov/18089281/)]
98. Myers TA, Crowther JH. Sociocultural pressures, thin-ideal internalization, self-objectification, and body dissatisfaction: could feminist beliefs be a moderating factor? *Body Image* 2007 Sep;4(3):296-308. [doi: [10.1016/j.bodyim.2007.04.001](https://doi.org/10.1016/j.bodyim.2007.04.001)] [Medline: [18089276](https://pubmed.ncbi.nlm.nih.gov/18089276/)]
99. Strahan EJ, Wilson AE, Cressman KE, Buote VM. Comparing to perfection: how cultural norms for appearance affect social comparisons and self-image. *Body Image* 2006 Sep;3(3):211-227. [doi: [10.1016/j.bodyim.2006.07.004](https://doi.org/10.1016/j.bodyim.2006.07.004)] [Medline: [18089224](https://pubmed.ncbi.nlm.nih.gov/18089224/)]
100. Trentowska M, Bender C, Tuschen-Caffier B. Mirror exposure in women with bulimic symptoms: how do thoughts and emotions change in body image treatment? *Behav Res Ther* 2013 Jan;51(1):1-6. [doi: [10.1016/j.brat.2012.03.012](https://doi.org/10.1016/j.brat.2012.03.012)] [Medline: [23168326](https://pubmed.ncbi.nlm.nih.gov/23168326/)]
101. Alleva JM, Sheeran P, Webb TL, Martijn C, Miles E. A meta-analytic review of stand-alone interventions to improve body image. *PLoS One* 2015 Sep 29;10(9):e0139177 [FREE Full text] [doi: [10.1371/journal.pone.0139177](https://doi.org/10.1371/journal.pone.0139177)] [Medline: [26418470](https://pubmed.ncbi.nlm.nih.gov/26418470/)]
102. Iper J, Sander C, Stein D. Pharmacotherapy and psychotherapy for body dysmorphic disorder. *Cochrane Database Syst Rev* 2009 Jan 21;2009(1):CD005332 [FREE Full text] [doi: [10.1002/14651858.CD005332.pub2](https://doi.org/10.1002/14651858.CD005332.pub2)] [Medline: [19160252](https://pubmed.ncbi.nlm.nih.gov/19160252/)]
103. Toole AM, Craighead LW. Brief self-compassion meditation training for body image distress in young adult women. *Body Image* 2016 Dec;19:104-112. [doi: [10.1016/j.bodyim.2016.09.001](https://doi.org/10.1016/j.bodyim.2016.09.001)] [Medline: [27664531](https://pubmed.ncbi.nlm.nih.gov/27664531/)]
104. Pliner P, Chaiken S, Flett GL. Gender differences in concern with body weight and physical appearance over the life span. *Pers Soc Psychol Bull* 1990 Jun 1;16(2):263-273. [doi: [10.1177/0146167290162007](https://doi.org/10.1177/0146167290162007)]
105. Gadalla TM. Eating disorders and associated psychiatric comorbidity in elderly Canadian women. *Arch Womens Ment Health* 2008 Dec;11(5-6):357-362. [doi: [10.1007/s00737-008-0031-8](https://doi.org/10.1007/s00737-008-0031-8)] [Medline: [18791783](https://pubmed.ncbi.nlm.nih.gov/18791783/)]

106. Lewis DM, Cachelin FM. Body image, body dissatisfaction, and eating attitudes in midlife and elderly women. *Eat Disord* 2001;9(1):29-39. [doi: [10.1080/106402601300187713](https://doi.org/10.1080/106402601300187713)] [Medline: [16864371](https://pubmed.ncbi.nlm.nih.gov/16864371/)]
107. McLaren L, Kuh D. Body dissatisfaction in midlife women. *J Women Aging* 2004;16(1-2):35-54. [doi: [10.1300/J074v16n01_04](https://doi.org/10.1300/J074v16n01_04)] [Medline: [15149923](https://pubmed.ncbi.nlm.nih.gov/15149923/)]
108. Reboussin BA, Rejeski WJ, Martin KA, Callahan K, Dunn AL, King AC, et al. Correlates of satisfaction with body function and body appearance in middle- and older aged adults: the activity counseling trial (ACT). *Psychol Health* 2000;15(2):239-254. [doi: [10.1080/08870440008400304](https://doi.org/10.1080/08870440008400304)]
109. Body image. National Eating Disorder Collaboration. URL: <https://nedc.com.au/eating-disorders/eating-disorders-explained/body-image/> [accessed 2021-08-27]
110. \$70 million for residential eating disorders treatment centres across Australia. Ministers Department of Health. URL: <https://www.health.gov.au/ministers/the-hon-greg-hunt-mp/media/70-million-for-residential-eating-disorders-treatment-centres-across-australia> [accessed 2021-08-27]
111. Milton AC, Hambleton A, Dowling M, Roberts AE, Davenport T, Hickie I. Technology-enabled reform in a nontraditional mental health service for eating disorders: participatory design study. *J Med Internet Res* 2021 Feb 16;23(2):e19532 [FREE Full text] [doi: [10.2196/19532](https://doi.org/10.2196/19532)] [Medline: [33591283](https://pubmed.ncbi.nlm.nih.gov/33591283/)]
112. Chrisler JC, Fung KT, Lopez AM, Gorman JA. Suffering by comparison: Twitter users' reactions to the Victoria's Secret Fashion Show. *Body Image* 2013 Sep;10(4):648-652. [doi: [10.1016/j.bodyim.2013.05.001](https://doi.org/10.1016/j.bodyim.2013.05.001)] [Medline: [23769929](https://pubmed.ncbi.nlm.nih.gov/23769929/)]
113. Ghaznavi J, Taylor LD. Bones, body parts, and sex appeal: an analysis of #thinspiration images on popular social media. *Body Image* 2015 Jun;14:54-61. [doi: [10.1016/j.bodyim.2015.03.006](https://doi.org/10.1016/j.bodyim.2015.03.006)] [Medline: [25880783](https://pubmed.ncbi.nlm.nih.gov/25880783/)]
114. Lewallen J, Behm-Morawitz E. Pinterest or Thinterest?: social comparison and body image on social media. *Soc Media Soc* 2016 Mar 30;2(1):205630511664055. [doi: [10.1177/2056305116640559](https://doi.org/10.1177/2056305116640559)]
115. Perloff RM. Social media effects on young women's body image concerns: theoretical perspectives and an agenda for research. *Sex Roles* 2014 May 29;71(11-12):363-377. [doi: [10.1007/s11199-014-0384-6](https://doi.org/10.1007/s11199-014-0384-6)]
116. Marengo D, Longobardi C, Fabris M, Settanni M. Highly-visual social media and internalizing symptoms in adolescence: the mediating role of body image concerns. *Comput Human Behav* 2018 May;82:63-69. [doi: [10.1016/j.chb.2018.01.003](https://doi.org/10.1016/j.chb.2018.01.003)]
117. Tigemann M, Slater A. NetGirls: the Internet, Facebook, and body image concern in adolescent girls. *Int J Eat Disord* 2013 Sep;46(6):630-633. [doi: [10.1002/eat.22141](https://doi.org/10.1002/eat.22141)] [Medline: [23712456](https://pubmed.ncbi.nlm.nih.gov/23712456/)]
118. Fardouly J, Diedrichs PC, Vartanian LR, Halliwell E. Social comparisons on social media: the impact of Facebook on young women's body image concerns and mood. *Body Image* 2015 Mar;13:38-45. [doi: [10.1016/j.bodyim.2014.12.002](https://doi.org/10.1016/j.bodyim.2014.12.002)] [Medline: [25615425](https://pubmed.ncbi.nlm.nih.gov/25615425/)]
119. Chen W, Wellman B. The global digital divide-within and between countries. *IT Society* 2004 Mar;1(7):39-45 [FREE Full text]
120. Milton AC, Ellis LA, Davenport TA, Burns JM, Hickie IB. Comparison of self-reported telephone interviewing and web-based survey responses: findings from the second Australian young and well national survey. *JMIR Ment Health* 2017 Sep 26;4(3):e37 [FREE Full text] [doi: [10.2196/mental.8222](https://doi.org/10.2196/mental.8222)] [Medline: [28951382](https://pubmed.ncbi.nlm.nih.gov/28951382/)]
121. Ethier RG, Poe GL, Schulze WD, Clark J. A comparison of hypothetical phone and mail contingent valuation responses for green-pricing electricity programs. *Land Econ* 2000 Feb;76(1):54-67. [doi: [10.2307/3147257](https://doi.org/10.2307/3147257)]
122. Burgess PM, Pirkis JE, Slade TN, Johnston AK, Meadows GN, Gunn JM. Service use for mental health problems: findings from the 2007 National Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatry* 2009 Jul;43(7):615-623. [doi: [10.1080/00048670902970858](https://doi.org/10.1080/00048670902970858)] [Medline: [19530018](https://pubmed.ncbi.nlm.nih.gov/19530018/)]
123. Nuttall F. Body mass index: obesity, BMI, and health: a critical review. *Nutr Today* 2015 May;50(3):117-128 [FREE Full text] [doi: [10.1097/NT.0000000000000092](https://doi.org/10.1097/NT.0000000000000092)] [Medline: [27340299](https://pubmed.ncbi.nlm.nih.gov/27340299/)]
124. Rothman KJ. BMI-related errors in the measurement of obesity. *Int J Obes (Lond)* 2008 Aug;32 Suppl 3:56-59. [doi: [10.1038/ijo.2008.87](https://doi.org/10.1038/ijo.2008.87)] [Medline: [18695655](https://pubmed.ncbi.nlm.nih.gov/18695655/)]
125. Taylor A, Dal Grande E, Gill T, Chittleborough CR, Wilson DH, Adams RJ, et al. How valid are self-reported height and weight? A comparison between CATI self-report and clinic measurements using a large cohort study. *Aust N Z J Public Health* 2006 Jun;30(3):238-246. [doi: [10.1111/j.1467-842x.2006.tb00864.x](https://doi.org/10.1111/j.1467-842x.2006.tb00864.x)] [Medline: [16800200](https://pubmed.ncbi.nlm.nih.gov/16800200/)]
126. Garner D. *Eating Disorder Inventory. Third Edition.* North Florida: PARInc; 2004.
127. Garner DM, Olmsted MP, Bohr Y, Garfinkel PE. The eating attitudes test: psychometric features and clinical correlates. *Psychol Med* 1982 Nov;12(4):871-878. [doi: [10.1017/s0033291700049163](https://doi.org/10.1017/s0033291700049163)] [Medline: [6961471](https://pubmed.ncbi.nlm.nih.gov/6961471/)]

Abbreviations

K10: 10-item Kessler Psychological Distress Scale

Edited by G Eysenbach; submitted 27.10.20; peer-reviewed by C Entwistle, S Six; comments to author 18.03.21; revised version received 13.05.21; accepted 31.05.21; published 04.11.21.

Please cite as:

Milton A, Hambleton A, Roberts A, Davenport T, Flego A, Burns J, Hickie I

Body Image Distress and Its Associations From an International Sample of Men and Women Across the Adult Life Span: Web-Based Survey Study

JMIR Form Res 2021;5(11):e25329

URL: <https://formative.jmir.org/2021/11/e25329>

doi: [10.2196/25329](https://doi.org/10.2196/25329)

PMID: [34734831](https://pubmed.ncbi.nlm.nih.gov/34734831/)

©Alyssa Milton, Ashlea Hambleton, Anna Roberts, Tracey Davenport, Anna Flego, Jane Burns, Ian Hickie. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 04.11.2021. 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

Role of 18F-Fluorodeoxyglucose–Positron Emission Tomography/Computed Tomography Imaging in the Prediction of Prognosis in Patients With Indolent Lymphoma: Prospective Study

Nawal Faiez AlShehry¹, MD; Raja Shanker^{2*}, MD; Syed Ziauddin Ahmed Zaidi^{1*}, MD; Fahad AlGhmlas², MD; Ibraheem Hussein Motabi¹, MD; Shahid Iqbal¹, MD; Ahmad Ali Butt¹, MD; Hassan AlShehri¹, MD; Imran Khan Tailor¹, MD; Syed Yasir Altaf¹, MD; Mubarak AlGhamdi¹, MD; Mohammed Marie¹, MD; Mansour AlFayez¹, MD; Kamal Al Zahrani¹, MD; Mohammed Dwaimah¹, MD; Tahani Al-Halouli³, MD; Wafaa Al-Shakweer³, MD; Maied Zaher AlShehry⁴, MD; Abdul Rehman Zia Zaidi¹, MD; Atta Munawar Gill¹, MD; Belal Mohammed Altoosh¹, BSc, RN; Musab Ahmed², MBBS

¹Department of Adult Hematology/Bone Marrow Transplantation, King Fahad Medical City, Riyadh, Saudi Arabia

²Department of Radiology and Nuclear Medicine, King Fahad Medical City, Riyadh, Saudi Arabia

³Department of Pathology and Laboratory Medicine, King Fahad Medical City, Riyadh, Saudi Arabia

⁴Department of Palliative Care, King Fahad Medical City, Riyadh, Saudi Arabia

*these authors contributed equally

Corresponding Author:

Nawal Faiez AlShehry, MD

Department of Adult Hematology/Bone Marrow Transplantation

King Fahad Medical City

Comprehensive Cancer Center

PO Box 59046

Riyadh, 11525

Saudi Arabia

Phone: 966 112889999 ext 23040

Email: nalshehry@kfmc.med.sa

Abstract

Background: The role of fluorodeoxyglucose–positron emission tomography/computed tomography (FDG-PET/CT) in indolent lymphoma has been minimally studied.

Objective: This study aims to assess the value of FDG-PET/CT in predicting the prognosis of indolent lymphoma.

Methods: We prospectively recruited 42 patients with indolent lymphoma. A total of 2 patients were excluded, and 40 underwent baseline PET/CT and follow-up at various time points. A total of 9 patients were observed only, 7 received 4 doses of rituximab alone, and 24 received chemoimmunotherapy. Metabolic response on follow-up PET/CT was assessed using the maximum standardized uptake value (SUVmax) and Deauville criteria (DC). We aimed to obtain the best SUVmax and DC to predict optimal survival rates, risk stratification, and optimize therapeutic strategies. The mean follow-up from the initial diagnosis was 33.83 months.

Results: SUVmax <4.35 at interim PET/CT provided the best discrimination, with a progression-free survival (PFS) of 100% and a median survival time of 106.67 months compared with SUVmax ≥4.35 ($P=.04$), which had a PFS of 43.8% and a median survival time of 50.17 months. This cutoff was also valuable in predicting overall survival at baseline, that is, 100% overall survival with baseline SUVmax <4.35, versus 58.4% for SUVmax ≥4.35 ($P=.13$). The overall survival of patients with a baseline DC score <3.0 was 100%, with a median overall survival of 106.67 months.

Conclusions: We demonstrated the utility of PET/CT in indolent lymphomas. SUVmax (<4.35 vs ≥4.35) on interim PET/CT performed best in predicting PFS.

(JMIR Form Res 2021;5(11):e24936) doi:[10.2196/24936](https://doi.org/10.2196/24936)

KEYWORDS

positron emission tomography; lymphoma; prognosis; indolent lymphoma; SUVmax; Deauville criteria

Introduction

Background

Indolent lymphomas are a heterogeneous group of mature B-cell non-Hodgkin lymphomas, characterized by a slow growth rate and a tendency to relapse. The frequencies of different types of indolent lymphomas based on the World Health Organization classification 2008 are follicular lymphoma (FL) (29%), small lymphocytic lymphoma (SLL) and chronic lymphocytic leukemia (CLL) (12%), mucosal-associated lymphoid tissue lymphoma (9%), nodal marginal zone lymphoma (2%), splenic marginal zone lymphoma (0.9%), and lymphoplasmacytic lymphoma (1.4%) [1]. The prognosis of indolent lymphoma varies according to the stage, clinical features, histology, immunophenotyping, response to first-line therapy, and the duration of remission [1]. Prognostic scores have been proposed specifically for the common types of FL [2,3], CLL [4], lymphoplasmacytic lymphoma [5], and splenic marginal zone lymphomas [6].

Indolent lymphomas mostly relapse to the same type of lymphoma or undergo histological transformation to more aggressive types of lymphomas. ¹⁸F-Fluorodeoxyglucose (FDG)-positron emission tomography (PET) has been valuable in evaluating the response to therapy, in addition to the diagnosis, staging, and surveillance of lymphoma patients. The maximum standardized uptake value (SUVmax) is a semiquantitative index used to objectively interpret the metabolic activity of various tissues (eg, neoplastic tissues) on PET/computed tomography (CT) scans. An international consensus helped establish a simpler, robust, and reproducible criterion for PET/CT interpretation in Hodgkin lymphoma (HL). The resultant 5-point Deauville criteria (DC) score helped in risk stratification, and strategizing adaptive therapies in HL, based on interim PET/CT findings with very high sensitivity and specificity. The DC established the role of PET/CT at the initial staging as well as at the end of the treatment [7,8]. Notably, the role of FDG-PET/CT in management has been scarcely studied internationally and is limited to FL and mantle cell lymphoma (MCL) [9,10].

The treatment of indolent lymphoma depends on whether the disease is localized or in an advanced stage. Therapeutic modalities range from watchful waiting to monoclonal anti-CD20 antibody therapy, conjugated-radiolabeled monoclonal antibody therapy, or more intensive chemoimmunotherapy and radiation therapy as well as high-dose chemotherapy and stem cell transplantation [11].

Superior tools for prognostication would help in risk stratification at the initial diagnosis or at the mid or end of treatment. High-risk patients, who are likely to relapse or undergo transformation to aggressive lymphomas, would be qualified to receive more intense therapeutic approaches and closer follow-up, particularly in the era of newly emerging targeted therapy for lymphomas. Emerging data suggest that metabolic response to FDG-PET/CT is superior in comparison

with other tools currently available, especially in early prognostication of HL and diffuse large B-cell lymphoma (DLBCL) [12-14]; the same may hold true in indolent lymphoma. In particular, the semiquantitative SUVmax has been found to objectively supplement visual interpretation of FDG-PET/CT [15], during and after therapy for HL and aggressive non-HL, and interim FDG-PET/CT was found to be an independent, strong predictor of progression-free survival (PFS) in HL [16]. Furthermore, in DLBCL, SUVmax has a high prognostic value and is strongly correlated with survival [12]. Similarly, SUVmax in DLBCL significantly improved prognostication after the 1st line of chemotherapy [12].

Although most indolent lymphomas show low metabolic activity of glucose, and hence low SUVmax, FDG-PET/CT has been found to be helpful in staging and evaluating histological transformation. Moreover, in FL, FDG avidity appears to be generally more pronounced than in other types of indolent lymphomas. PET/CT is also more accurate than conventional imaging for initial staging, often prompting significant management changes [17].

The results of a clinical trial [17] found that FDG-PET/CT should be considered as a new standard for response assessment of FL and guiding response-adapted therapy. In another study, end-of-treatment PET/CT findings were found to be predictive of overall survival in patients with FL [18]. A review by El-Galaly et al [19] highlighted the role of PET/CT in staging and its impact on modern treatment selection for lymphomas, including the indolent group of lymphomas.

Objective

We prospectively studied the utility of FDG-PET/CT in the management of indolent lymphomas at our tertiary care institution in Saudi Arabia. This study aims to assess the value of FDG-PET/CT in predicting the prognosis of indolent lymphomas.

Methods

Study Design

This was a prospective study conducted at King Fahad Medical City, Riyadh, Saudi Arabia, after approval by the Institutional Review Board under study number 15-262.

Patient Population and Sampling Technique

All newly diagnosed adult patients with different types of indolent lymphomas were enrolled in the study after obtaining informed consent from August 2014 to May 2018. A 1-year follow-up scan for the last recruited patient was completed in May 2019. Data analysis was completed in February 2020.

Inclusion Criteria

All patients aged ≥ 14 years (as patients in this age group are treated by the adult team in our hospital), a confirmed diagnosis of indolent lymphoma (other than FL grade 3b), any stage (Ann Arbor I, II, III, IV and in CLL, Rai stage 0-IV), and

those who underwent a PET/CT scan before treatment or start of observation were included in the study.

Exclusion Criteria

Patients with grade 3b FL were excluded. Patients could withdraw from the study for any reason.

Treatment Guidelines

On the basis of the initial histology and stage, the treatment options were as follows:

- Chemotherapy+Immunotherapy (chemoimmunotherapy)
- Rituximab only
- Observation only

All these factors were considered for finding any significant correlations between the PET/CT scan and the specified endpoints.

Primary Objective

To assess the utility of FDG-PET/CT in prognostication (based on SUVmax and DC) at various time points (baseline, midtreatment, end of treatment, and 1-year follow-up).

Secondary Objectives

- To evaluate the role of FDG-PET/CT in adaptive therapies, that is, intensive chemotherapy or targeted therapy.
- To evaluate the utility of PET/CT in situations such as disease progression or transformation of indolent lymphoma to DLBCL.

Data Collection Procedures

The patients were managed according to current standards of practice. Case report forms were completed for all patients with data entry at various phases of the study. The clinical data were collected from electronic medical records.

All newly diagnosed patients (proven by histopathology) had pretreatment baseline evaluation with FDG-PET/CT and follow-up PET/CT, that is, interim PET/CT, end-of-treatment PET/CT, and PET/CT 1 year later. FDG uptake was measured by SUVmax normalization for total body weight, in addition to the DC. The DC is a five-point scale that uses the mediastinum and liver as reference organs. The patients were followed

prospectively to assess the prognostic value of early and end-of-treatment PET/CT in these patients.

The primary endpoints of this study were the determination of PFS in correlation with PET/CT findings (SUVmax and DC) at different time points.

FDG-PET/CT Procedure

All PET/CT scans were acquired per accepted protocol following a tracer dose of 8-12 mCi of FDG for adults, or weight-adjusted dosage in smaller patients on the same GE 960 STE scanner (General Electric, Waukesha, WI, USA). All scans were performed using the same image acquisition and reconstruction protocol. PET/CT was performed from the vertex of the skull to the knees 60 minutes after injection and took an average time of 30 min. Both the uncorrected and attenuation-corrected images were reviewed. All PET/CT scans were reviewed visually and semiquantitatively (using the SUVmax and DC) and interpreted blindly by 2 experienced nuclear medicine physicians. The PET/CT images and interpretations were reviewed in clinical context by experienced hematologists.

Statistical Analysis

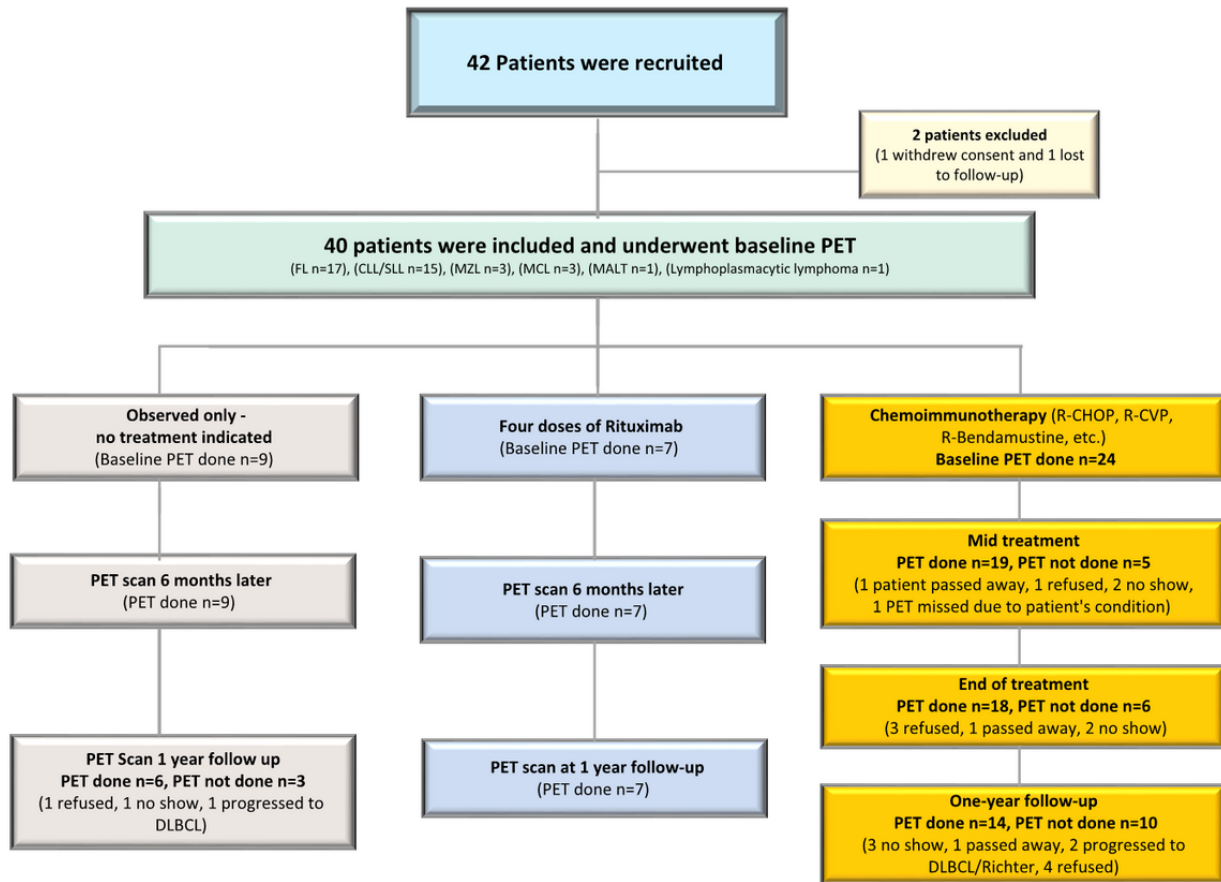
All the data were analyzed using SPSS, version 25 (IBM Corporation). The primary endpoints and outcomes, with complete follow-up assessment of the study, were evaluated by the difference in response rates between patients treated with chemoimmunotherapy and rituximab alone. The log-rank test was applied to determine the time survival between 2 or more independent groups. Kaplan–Meier (KM) survival curves were used to estimate the PFS and overall survival of the patients. Relapse-free survival (RFS) was also estimated for patient groups treated with rituximab and chemoimmunotherapy. $P < .05$ was considered as statistically significant.

Results

Overview

Of the 42 patients initially enrolled in the study, 40 were analyzed. Two patients were excluded: one due to withdrawal of consent and the other was lost to follow-up (Figure 1).

Figure 1. Patient allocation and schema of follow-up during the study. CLL: chronic lymphocytic leukemia; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; MALT: mucosa-associated lymphoid tissue lymphoma; MCL: mantle cell lymphoma; MZL: marginal zone lymphoma; PET: positron emission tomography; SLL: small lymphocytic lymphoma; R-CHOP: rituximab–cyclophosphamide, doxorubicin hydrochloride, Oncovin, prednisone; R-CVP: rituximab–cyclophosphamide, vincristine sulfate, prednisone.



Patient Characteristics

The study population included 68% (27/40) men and 33% (13/40) women. Patients’ ages ranged from 28 to 88 years with a mean of 59.10 (SD 14.78). The diagnoses included FL (17/40,

43%), SLL and CLL (15/40, 38%), marginal zone lymphoma (3/40, 8%), MCL (3/40, 8%), mucosal-associated lymphoid tissue lymphoma (1/40, 3%), and lymphoplasmacytic lymphoma (1/40, 3%; Table 1).

Table 1. Basic demography and clinical characteristics of the patients (N=40).

Characteristics	Participants
Age at 1st diagnosis (years) ^a , mean (SD)	59.10 (14.78)
Gender, n (%)	
Male	27 (68)
Female	13 (33)
Therapeutic strategy, n (%)	
Chemoimmunotherapy	24 (60)
Rituximab	7 (18)
Observation	9 (23)
Diagnosis, n (%)	
Follicular	17 (43)
SLL ^b and CLL ^c	15 (38)
Marginal zone lymphoma	3 (8)
Mantle cell	3 (8)
MALT ^d	1 (3)
Lymphoplasmacytic lymphoma	1 (3)
Lymphadenopathy, n (%)	
Yes	31 (78)
No	9 (23)
Organomegaly, n (%)	
Yes	14 (35)
No	26 (65)
Bone marrow infiltration at diagnosis, n (%)	
Yes	28 (70)
No	9 (23)
Not done	3 (8)
Bulky disease, n (%)	
Yes	5 (13)
No	35 (88)
Not done	0 (0)
Final staging, n (%)	
Early stage	11 (28)
Advanced-stage disease	29 (73)
IPI^e, n (%)	
Not applicable	38 (95)
Intermediate	1 (35)
High	1 (3)
FLIPI^f, n (%)	
Not applicable	24 (60)
Low	1 (35)
Intermediate	3 (8)
High	12 (30)

Characteristics	Participants
Rituximab with 1st line, n (%)	
Yes	31 (78)
No	3 (8)
Not applicable	6 (15)
Rituximab maintenance, n (%)	
Yes	9 (23)
No	22 (55)
Not applicable	9 (23)
Relapse status, n (%)	
Yes	5 (13)
No	29 (73)
Outcome, n (%)	
Died	5 (13)
Alive	34 (85)

^aThe median age of participants is 61 years (range 28-82 years).

^bSLL: small lymphocytic lymphoma.

^cCLL: chronic lymphocytic leukemia.

^dMALT: mucosal-associated lymphoid tissue.

^eIPI: international prognostic index.

^fFLIPI: follicular lymphoma international prognostic index.

All 40 patients underwent baseline PET/CT. On the basis of clinical presentation, the patients were managed per accepted protocol as follows: either observed only (n=9), received 4 doses of rituximab alone (n=7), or received chemoimmunotherapy (R-CHOP [rituximab-cyclophosphamide, hydroxydaunorubicin, vincristine and prednisone], R-CVP [cyclophosphamide, vincristine and prednisone], and R-bendamustine; n=24). Most of the patients underwent

scheduled FDG-PET/CT scans at midtreatment (only chemoimmunotherapy arm), at 6 months, and at 1 year. However, due to logistics, we could not obtain PET/CT scans in a few patients in the later phases on 16 occasions (Figure 1). For the overall study population, SUVmax and DC scores at baseline, interim, and end of treatment per 1 year observation as well as 1 year after chemotherapy are summarized in Table 2.

Table 2. Descriptive analysis of the study population for maximum standardized uptake value (SUVmax) and Deauville criteria.

Characteristics	Value, median (range)	Value, mean (SD)
SUVmax baseline	6 (2-24)	8 (5)
SUVmax interim	4 (1-25)	5 (4)
SUVmax Eed of treatment per 1 year observation	3 (1-25)	5 (5)
SUVmax at 1 year after last C or chemotherapy	4 (2-14)	5 (3)
Deauville baseline	5 (2-5)	4 (1)
Deauville interim	3 (1-5)	3 (1)
Deauville end treatment per 1 year	2 (1-5)	2 (1)
Deauville at 1 year after last chemoimmunotherapy	2 (1-5)	3 (1)

The mean time for follow-up of chemoimmunotherapy patients was 33.24 months (SD 18.2); whereas it was 35.99 months (SD 11.65) and 38.83 months (SD 30.82) for rituximab and observation groups, respectively.

Due to the limited sample sizes of different subsets of indolent lymphomas, we chose to compare patients with FL (n=17) against a subset of patients with nonfollicular indolent lymphoma (n=23) as presented in Table 3. Bulky disease was

seen in 29% (5/17) patients with FL, whereas 71% (12/17) of patients with FL had high follicular lymphoma international prognostic index (FLIPI) scores. Of the 40 patients, 73% (29/40) had advanced-stage disease, more so in the FL group with 88% (15/17) patients than in the nonfollicular indolent lymphoma group, with 61% (14/23; $P=.06$) patients. Lymphadenopathy was observed in 94% (16/17) patients with FL, compared with 65% (15/23) in the nonfollicular indolent lymphoma patients

($P=.03$). However, organomegaly was more common in the non-FL group ($P=.05$) as presented in Table 3.

Table 3. Relationship between the diagnosis of follicular lymphoma (FL) versus non-FL type of indolent lymphomas and characteristics of patients (N=40).

Characteristics	Follicular (n=17), n (%)	Nonfollicular (n=23), n (%)	P value
Gender			.27
Male	12 (71)	15 (65)	
Female	5 (29)	8 (35)	
B symptoms			.68
Yes	7 (41)	8 (35)	
No	10 (59)	15 (65)	
Lymphadenopathy			.03
Yes	16 (94)	15 (65)	
No	1 (6)	8 (35)	
Organomegaly			.048
Yes	3 (18)	11 (48)	
No	14 (82)	12 (52)	
Bone marrow infiltration at diagnosis			.24
Yes	12 (71)	16 (70)	
No	5 (29)	4 (17)	
Not done	0 (0)	3 (13)	
Bulky disease			.005
Yes	5 (29)	0 (0)	
No	12 (71)	23 (100)	
Final staging			.06
Early stage	2 (12)	9 (39)	
Advanced-stage disease	15 (88)	14 (61)	
IPI^a			.50
Not applicable	17 (100)	21 (91)	
Intermediate	0 (0)	1 (4)	
High	0 (0)	1 (4)	
FLIPI^b			.001
Not applicable	1 (6)	23 (100)	
Low	1 (6)	0 (0)	
Intermediate	3 (18)	0 (0)	
High	12 (71)	0 (0)	

^aIPI: international prognostic index.

^bFLIPI: follicular lymphoma international prognostic index.

As expected, the advanced stage of indolent lymphoma had a higher SUVmax in comparison with early-stage indolent lymphomas (Table 4). Patients with a DC score >3.0, at baseline, had higher lactate dehydrogenase compared with those with a DC score ≤3.0 (20/23, 87% vs 3/23, 13%; Table 5). Patients with FL had higher SUVmax and DC scores than all other nonfollicular indolent lymphomas (Table 6). Twelve patients

with FL had bone marrow (BM) involvement on biopsy, but in one of these patients, BM involvement could not be detected on PET/CT (Figure 2).

Patients with a Deauville score >3.0 at baseline had high lactate dehydrogenase (20/23, 87%) compared with those with a Deauville score ≤3.0 (3/23, 13%).

Table 4. Distribution of patients according to the maximum standardized uptake value (SUVmax), Deauville and final staging (N=40).

Characteristics	Disease staging, n (%)		P value
	Early stage	Advanced-stage disease	
Group SUVmax baseline			<.001
<4.35	9 (82)	4 (14)	
≥4.35	2 (18)	25 (86)	
Group SUVmax baseline			.007
<9.41	11(100)	16 (55)	
≥9.41	0 (0)	13 (45)	
Group Deauville baseline			<.001
≤3.0	8 (73)	2 (7)	
>3.0	3 (27)	27 (93)	
Group Deauville baseline			<.001
≤4.0	10 (91)	7 (24)	
>4.0	1 (9)	22 (76)	

Table 5. Distribution of patients according to the maximum standardized uptake value (SUVmax), Deauville score, and lactate dehydrogenase (N=40).

Characteristics	LDH ^a , n (%)		P value
	≤220 (normal; n=17)	>220 (high; n=23)	
SUVmax baseline			.31
<4.35	7 (41)	6 (26)	
≥4.35	10 (59)	17 (74)	
SUVmax baseline			.29
<9.41	13 (76)	14 (61)	
≥9.41	4 (24)	9 (39)	
Deauville baseline			.04
≤3.0	7 (41)	3 (13)	
>3.0	10 (59)	20 (87)	
Deauville baseline			.25
≤4.0	9 (53)	8 (35)	
>4.0	8 (47)	15 (65)	

^aLDH: lactate dehydrogenase.

Table 6. Distribution of patients according to the maximum standardized uptake value (SUVmax), Deauville, and diagnosis (N=40).

Characteristics	Diagnosis, n (%)		P value
	Follicular	Nonfollicular	
Group SUVmax baseline			.002
<4.35	1 (6)	12 (52)	
≥4.35	16 (94)	11 (48)	
Group SUVmax baseline			<.001
<9.41	6 (35)	21 (91)	
≥9.41	11 (65)	2 (9)	
Group Deauville baseline			.02
≤3.0	1 (6)	9 (39)	
>3.0	16 (94)	14 (61)	
Group Deauville baseline			.006
≤4.0	3 (18)	14 (61)	
>4.0	14 (82)	9 (39)	

Figure 2. Sequential fluorodeoxyglucose–positron emission tomography/computed tomography scan findings in subject number 2, a case of follicular lymphoma.

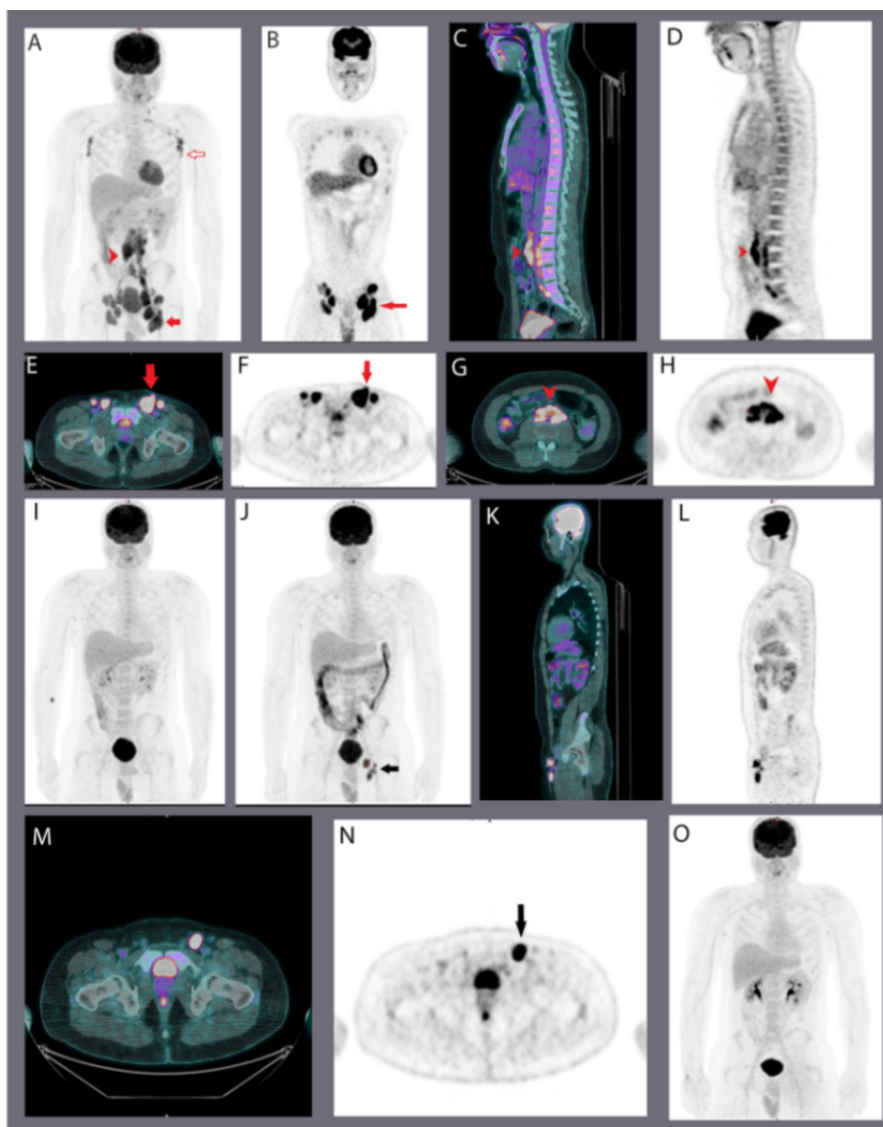
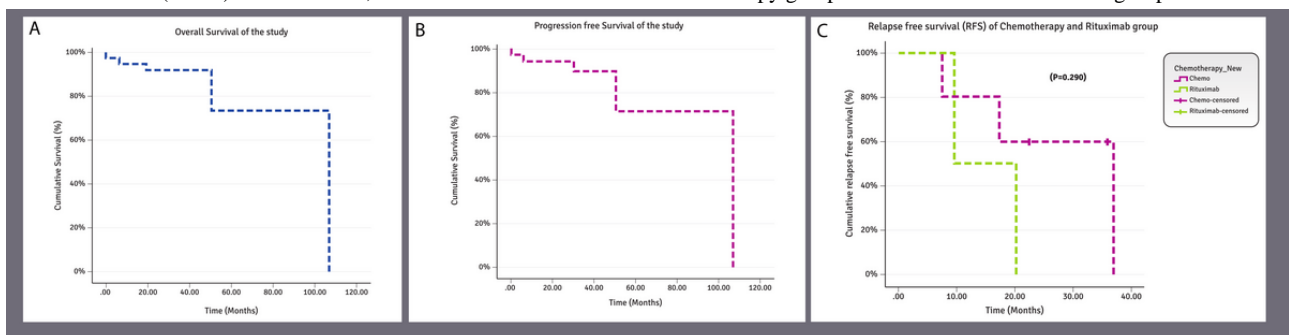


Figure 2 shows the images of the patient from initial presentation (images A-H) until complete metabolic remission. Maximum intensity projection (MIP; image A) and multiplanar PET/CT demonstrated extensive hypermetabolic lymphadenopathy above and below the diaphragm. The arrowhead indicates retroperitoneal lymphadenopathy, and the solid arrow indicates left inguinal lymphadenopathy (A, MIP; B, coronal PET; C, sagittal fused PET/CT; D, sagittal PET; E, axial fused PET/CT; F, axial PET; G, axial fused PET/CT; and H, axial PET). BM involvement was not observed. The patient achieved complete metabolic remission of previously noted residual hypermetabolic lymphadenopathy above and below the diaphragm after 6 cycles of R-CVP and rituximab maintenance (image I, MIP). He

relapsed within 21 months with interval development of FDG-avid left inguinal lymphadenopathy (images J-N). Solid arrow: left inguinal lymphadenopathy (J, MIP; K, sagittal fused PET/CT; L, sagittal PET; M, axial fused PET/CT; and N, axial PET). After involved-field radiation to the inguinal area, the patient achieved complete metabolic remission of previously noted left inguinal lymphadenopathy (image O, MIP).

The cumulative overall survival of the study participants was 73.4%. The estimated mean overall survival time was 88.29 months (range 66.28-110.31 months). The cumulative PFS of the study sample was 71.7%, with a median PFS time of 87.14 months (range 64.19-109.37; Figure 3).

Figure 3. (A) Overall survival of all the study patients (N=40). The cumulative overall survival rate of the study was 73.4%. However, it was not possible to obtain the median survival time of this curve, as the censorship was less than 50% but the estimated overall mean survival time was recorded as 88.29 (range 66.28–110.31) months. (B) Progression-free survival (PFS) of all the study patients (N=40). The overall cumulative PFS rate of the study was 71.7% and the overall median progression survival time was 87.14 (range 64.91–109.37) months. (C) Comparison of the cumulative RFS of the patients treated with chemoimmunotherapy (n=24) or rituximab only (n=7). The cumulative RFS of the patients treated with chemoimmunotherapy was 60% and the median relapse-free survival time was observed to be 36.60 months, whereas the cumulative RFS of the patients treated with rituximab only was 50% and the median relapse-free survival time was observed to be 9.67 months. No statistically significant difference was observed between the 2 survival curves (P=.29). At 15 months, RFS was 80% for the chemoimmunotherapy group versus 50% for the rituximab group.

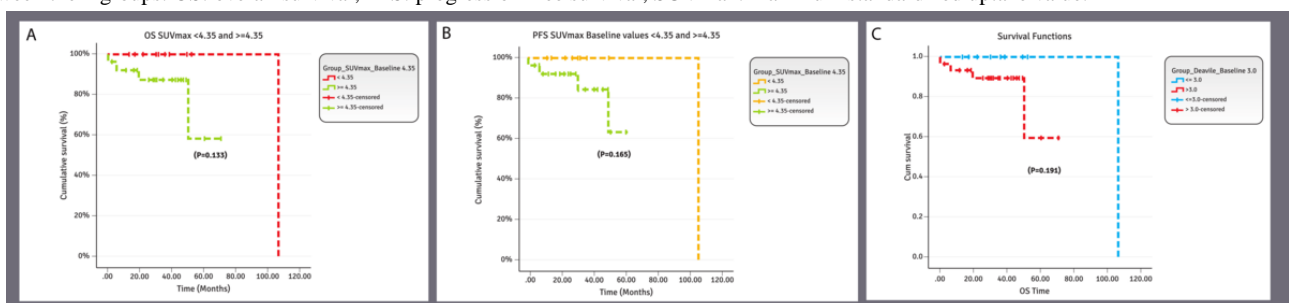


Survival curves were developed (overall survival and PFS) for all study participants (Figures 3A and B), and a comparative overall survival and PFS KM curve based on various SUVmax values was developed to find the best SUVmax cutoff values that might guide the estimation of survival.

These included survival curves for the baseline PET/CT scan SUVmax ≥ 4.35 vs < 4.35 , and ≥ 9.41 vs < 9.41 (33rd and 66th

percentile-based values). Using the receiver operating characteristic curve, we were unable to detect an optimum threshold for SUVmax at baseline to predict overall survival; however, we adjudicated that SUVmax value of 4.35 (33rd percentile-based values) was a good cutoff value for segregating survival function at baseline (Figure 4).

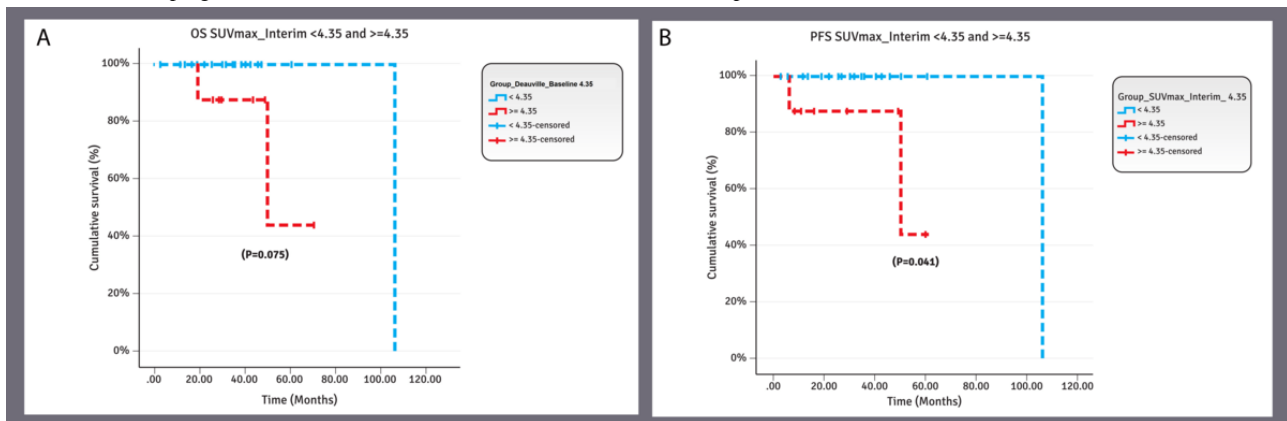
Figure 4. (A) Overall survival maximum SUVmax baseline values < 4.35 and ≥ 4.35 . The overall cumulative survival rate of the patients with baseline SUVmax < 4.35 was 100% and the median survival time was observed to be 106.67 months, whereas the cumulative survival of the patients with baseline SUVmax ≥ 4.35 was 58.4% and the median survival time cannot be calculated because the censorship is less than 50%. Moreover, there were no statistically significant differences between the 2 groups (P=.13). (B) PFS SUVmax baseline values < 4.35 and ≥ 4.35 . The cumulative PFS of the patients with SUVmax baseline < 4.35 was 100% and the median survival time was 106.67 months, whereas the cumulative survival of the patients with SUVmax baseline ≥ 4.35 was 63% and no median survival time was observed. Furthermore, there was no statistically significant difference between the 2 groups (P=.17). (C) Overall survival Deauville baseline values ≤ 3.0 and > 3.0 . The overall cumulative survival of the patients with baseline Deauville < 3.0 was 100% and the median survival time (months) was 106.67, whereas the overall survival of the patients with baseline Deauville ≥ 3.0 was 59.30% and the median survival time cannot be calculated because the censorship is less than 50%. Moreover, there were no statistically significant differences (P=.19) between the 2 groups. OS: overall survival; PFS: progression-free survival; SUVmax: maximum standardized uptake value.



For follow-up phases (after treatment per 6-month follow-up and at 1-year follow-up), we compared overall survival with PFS for SUVmax ≥ 4.35 vs < 4.35 (Figures 5A and B). Comparative overall survival and PFS KM curves for Deauville

score ≤ 3 vs < 3 were also developed at presentation (Figure 4C), after treatment per 6-month follow-up, and at 1-year follow-up (not shown).

Figure 5. (A) OS SUVmax interim values < 4.35 and ≥ 4.35 . The overall cumulative survival of the patients with SUVmax interim < 4.35 was 100% and the median survival time was observed to be 106.67 months, whereas the cumulative survival of the patients with SUVmax interim ≥ 4.35 was 43.8% and the median survival time was 50.37 months. No statistically significant difference was observed between the 2 survival curves ($P=.08$). (B) PFS SUVmax interim values < 4.35 and ≥ 4.35 . The cumulative PFS of the patients with SUVmax interim < 4.35 was 100% and the median survival time was observed to be 106.67 months, whereas the cumulative survival of the patients with SUVmax interim ≥ 4.35 was 43.8% and the median survival time was observed to be 50.170 months. Furthermore, statistically significant differences were observed between the 2 survival curves ($P=.04$). OS: overall survival; PFS: progression-free survival; SUVmax: maximum standardized uptake value.



The Best Discriminatory SUVmax Values of This Study

Overall, the value of SUVmax (< 4.35 vs ≥ 4.35) on interim PET/CT performed the best discriminatory function in predicting PFS. The cumulative PFS of the patients with SUVmax < 4.35 at interim PET/CT was 100% and the observed median survival time was 106.67 months. Whereas the cumulative PFS of the patients with SUVmax ≥ 4.35 on interim PET/CT was 43.8% and the observed median survival time was 50.17 months. Furthermore, a statistically significant difference was observed between the 2 survival curves ($P=.04$; Figure 5B).

The same value of SUVmax (< 4.35 vs ≥ 4.35) showed good discrimination in predicting overall survival at baseline and interim PET/CT. The cumulative overall survival of the patients with baseline SUVmax < 4.35 was 100% and the median survival time was 106.67 months; whereas the cumulative overall survival of the patients with baseline SUVmax ≥ 4.35 was 58.4% ($P=.13$; Figure 4A). The cumulative overall survival of the patients with an interim SUVmax < 4.35 was 100% and the median survival time was observed to be 106.67 months. Whereas the cumulative overall survival of the patients with interim SUVmax ≥ 4.35 was 43.8% and the median survival time was 50.37 months. Although there was no statistically

significant difference between the 2 groups ($P=.08$), there was a trend of better overall survival in the group with interim SUVmax < 4.35 (Figure 5A).

The cumulative PFS of the patients with SUVmax baseline < 4.35 was 100% and the median survival time was 106.67 months, whereas the cumulative PFS of the patients with baseline SUVmax ≥ 4.35 was 63%. This difference was not statistically significant ($P=.17$; Figure 4B).

Indolent lymphomas, due to their nature, are deemed to relapse; thus, we did not focus on a detailed analysis of RFS. However, we compared the cumulative RFS of the patients treated with chemoimmunotherapy ($n=24$) and those of patients treated with rituximab only ($n=7$). The cumulative RFS of the chemoimmunotherapy group was 60% and the median RFS time was 36.6 months, whereas RFS of the *rituximab-only* group was 50% with a median RFS time of 9.67 months ($P=.29$; Figure 3C).

In our study, 3 patients underwent transformation to DLBCL during the study period (eg, Figures 6 and 7). The SUVmax at transformation was 11, 14.3, and 19.8, respectively. This was helpful in guiding targeted biopsies in patients with clinically suspected transformation.

Figure 6. Sequential fluorodeoxyglucose–positron emission tomography/computed tomography (FDG-PET/CT) scan findings in participant number 14, a case of small lymphocytic lymphoma (SLL) and chronic lymphocytic leukemia (CLL) that transformed to diffuse large B-cell lymphoma. Initial presentation showed multiple low FDG-avid lymphadenopathy in the head and neck, mesenteric, bilateral axillary, and bilateral inguinal regions (images A-D). Solid arrows: bilateral axillary lymphadenopathy (A, maximum intensity projection [MIP]; B, coronal fused PET/CT; C, axial fused PET/CT; and D, axial PET). On follow up, interval development of new intense FDG-avid mesenteric lymph nodes (arrowhead in images E-H) and biopsy showed persistent SLL with Ki67 of 10% (E, coronal fused PET/CT; F, MIP; G, axial fused PET/CT; and H, axial PET). The patient sought a second opinion and was lost to follow-up until he reported with interval development of new intense FDG-avid large right subpectoral lymph node and interval resolution of the FDG-avid mesenteric lymph node (images I-K). Solid arrows: right subpectoral lymphadenopathy (I, MIP; J, axial fused PET/CT; and K, axial PET/CT). After R-mini-CHOP (cyclophosphamide, hydroxydaunorubicin, vincristine, and prednisone) chemotherapy, there was complete metabolic response of the previously noted intense FDG-avid subpectoral lymph nodes as well as amild FDG activity in the bilateral axillary lymph nodes (L, MIP).

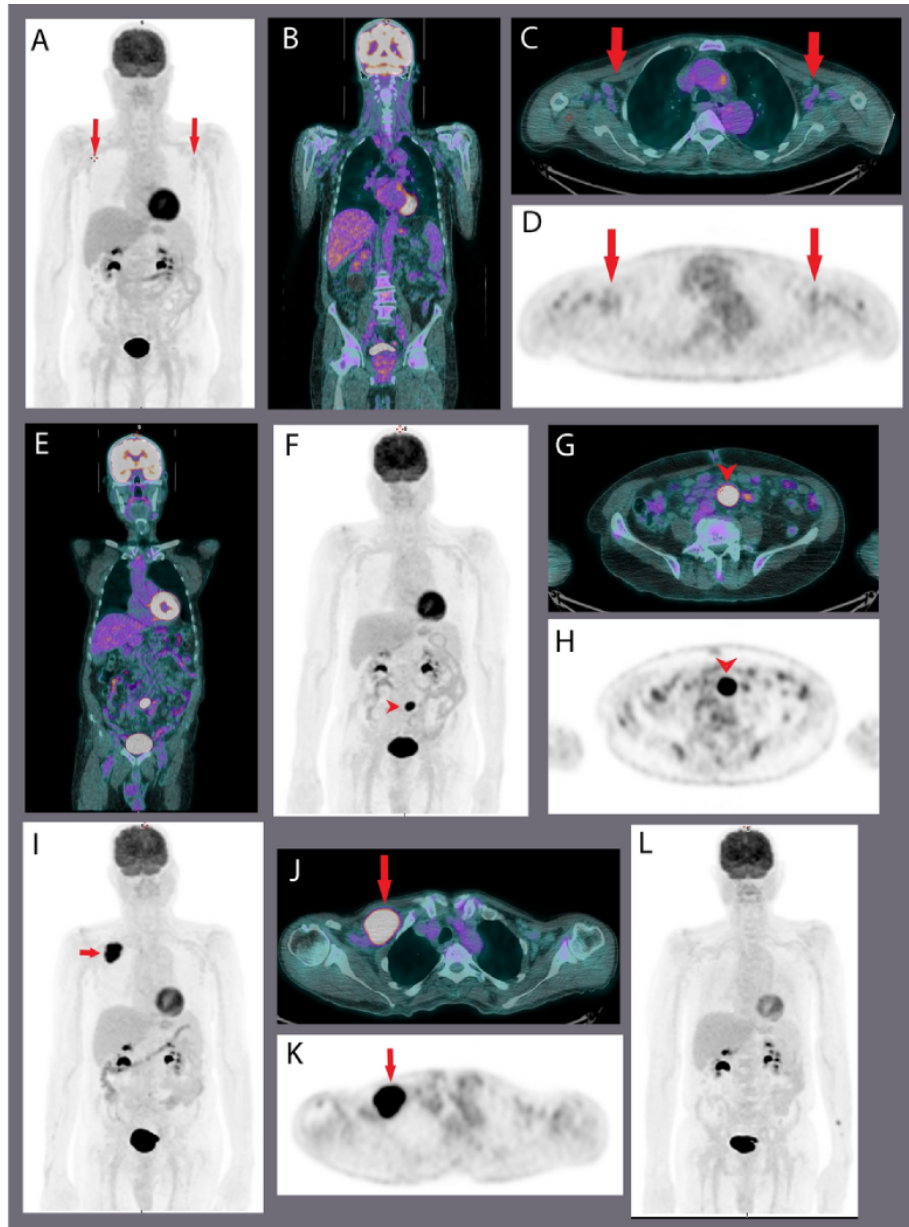
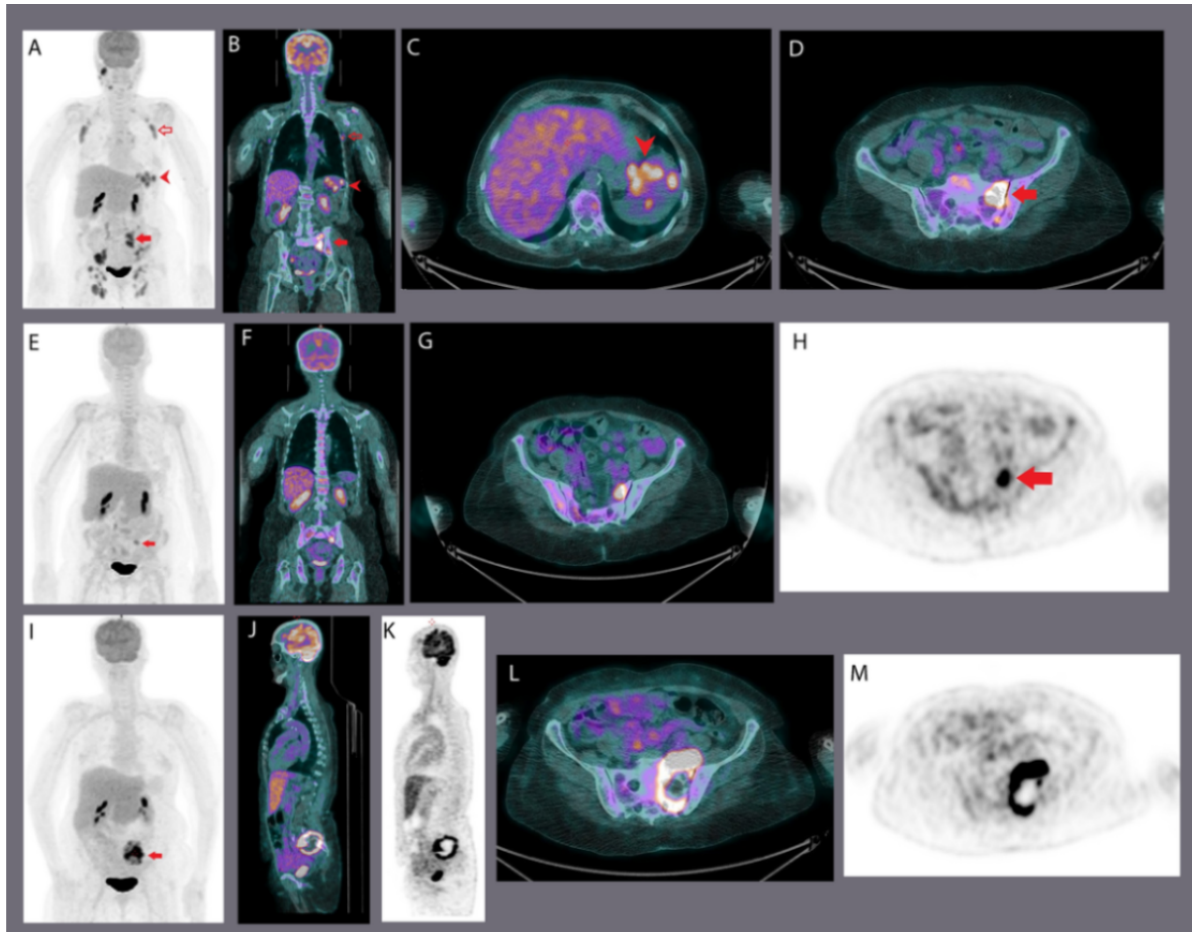


Figure 7. Sequential fluorodeoxyglucose–positron emission tomography/computed tomography (FDG-PET/CT) scan findings in participant number 40, a case of follicular lymphoma that transformed to diffuse large B-cell lymphoma (DLBCL). At baseline, follicular lymphoma causing multiple FDG-avid lymphadenopathy was observed above and below the diaphragm with splenic and osseous involvement (images A-D). Solid arrow: FDG-avid left sacral ala bone metastasis; hollow arrow: FDG-avid left axillary lymphadenopathy; arrowhead: multiple focal areas of FDG uptake in the spleen (A, maximum intensity projection [MIP]; B, coronal fused PET/CT; C, axial fused PET/CT; and D, axial fused PET/CT). After 6 cycles of rituximab + bendamustine, interval metabolic resolution (images E-H) of the previously noted FDG-avid lymphadenopathy and splenic lesion and most of the bony lesions except residual FDG-avid solitary bone lesion was noted in the sacrum (arrows: E, MIP; F, coronal fused PET/CT; G, axial fused PET/CT; and H, axial PET). On further follow-up, an interval increase in size and FDG activity of the left sacral ala metastasis (arrow) was proven to be transformation to DLBCL (images I-M) (I, MIP; J, sagittal fused PET/CT; K, sagittal PET; L, axial fused PET/CT; M, axial PET).



Discussion

Principal Findings

In this prospective study evaluating the utility of PET/CT in indolent lymphomas, FDG uptake was measured using the SUVmax and Deauville criteria at various time points. We developed the overall survival and PFS curves for the study population ($n=40$) and comparative overall survival and PFS KM curves based on SUVmax and Deauville score at baseline and various follow-up phases. Moreover, we evaluated the role of FDG-PET/CT in adaptive therapies, that is, intensive chemotherapy and/or targeted therapy, as well as in situations such as transformation or progression of indolent lymphomas to DLBCL. PET/CT was helpful in detecting SLL and CLL transformation to DLBCL (Figure 6), and FL relapse or transformation to DLBCL (Figure 7).

Median RFS for the rituximab-only group was 50% compared with 60% for the chemoimmunotherapy group, and the median RFS time was 9.67 months and 36.6 months respectively ($P=.29$), indicating insignificant superiority of

chemoimmunotherapy. The outcome in patients with FL has improved over the last 2 decades through the introduction of anti-CD20 monoclonal antibodies, which are usually used in combination with chemotherapy. Chemotherapy-free rituximab only has been proposed as a preferable approach in low tumor burden FL, but is still a matter of debate [20-22]. Lockmer et al [23] indicated that an initial rituximab-only approach in patients with indolent lymphomas was associated with an increased overall survival compared with that found in other studies with first-line immunochemotherapy. However, our study, although having a short follow-up time, showed better outcomes with chemoimmunotherapy than with rituximab only.

The most significant findings included an SUVmax cutoff value of 4.35 at interim assessment. The cumulative PFS of the patients with SUVmax interim <4.35 was 100%, and the observed median survival time was 106.67 months as opposed to the patients with SUVmax interim ≥ 4.35 who had a PFS of 43.8% and a median survival time of 50.17 months ($P=.04$). Patients with interim SUVmax <4.35 had an overall survival of 100% with a median survival time of 106.67 months. For patients with interim SUVmax ≥ 4.35 , overall survival was

43.8% with a median survival time of 50.37 months ($P=.08$). Regarding the utility of the Deauville score in indolent lymphomas, we noted a mild trend: at baseline, the overall survival was observed to be better with Deauville score <3.0 compared with Deauville score ≥ 3.0 . The cumulative overall survival of the patients with baseline Deauville <3.0 was 100%, whereas the overall survival of the patients with baseline Deauville ≥ 3.0 was 59.30% ($P=.19$). In addition, it confirmed the utility of a remarkably high SUVmax in PET/CT for the transformation of indolent lymphomas in 3 patients.

It can be speculated that a relatively short follow-up period of a small population size divided among 3 management strategies contributed to the marginal degree of segregation of survival curves based on the chosen cutoff values for SUVmax we studied. The median follow-up time of our study population from diagnosis was 33.50 months (mean 34.83 months). The follow-up duration for chemoimmunotherapy patients was 33.24 months (SD 18.2), for the rituximab group it was 35.99 months (SD 11.65), and for the observed group it was 38.83 months (SD 30.82). As indolent lymphomas have a characteristic behavior of being incurable with commonly available therapies, and relapse of the same type of lymphoma or histological transformation to more aggressive types of lymphomas is common, this follow-up period is not sufficient to draw more reasonable results. Expanding the follow-up, as well as increasing the sample size would have yielded statistically robust values.

FDG may be used as a biomarker for assessment of response to therapy by either visual interpretation, such as the DC or by semiquantitative assessment of response using an SUVmax threshold. Currently, ^{18}F -FDG-PET/CT is considered the standard imaging technique for evaluating response assessment in HL, DLBCL, and FL due to its ability to reveal residual metabolic activity irrespective of residual volume. The role of baseline SUVmax in predicting therapeutic outcomes in indolent lymphoma has also been investigated [8]. In our study, SUVmax in FL was 2.1-24 (mean 8). For patients with FL with a high FLIPI score, the SUVmax was 6.1-24 (mean 10.4). All patients with FL, except 1, were alive at the time of the analysis. PFS in FL cases was 90-576 days (mean 780.2 days). For those with high FLIPI, the PFS was 120-1823 (mean 787.3). The overall survival time in all FL cases was 90-1823 (mean, 944.8). In those with a high FLIPI score, the overall survival time was 355-1823 (mean 941.8). Hence, we did not observe the impact of SUVmax on the survival outcome of patients with FL. In concordance with the findings of Trotman et al [17,24], our study, although limited by the number of FL cases (17, of which 5 had bulky disease and 12 had high FLIPI score), we did not observe any correlation between FLIPI score and outcome.

A retrospective study of 81 patients with MCL revealed that baseline SUVmax is predictive of outcome; in patients with SUVmax <5.0 , the 5-year survival was 87.7%, whereas it dropped to 34% when SUVmax was >5.0 [25]. In our study, 3 patients had MCL, and their baseline SUVmax ranged from 5.2-6.8. The overall survival time ranged from 181 to 1379 days (median 1051). Although all 3 had advanced-stage disease and received similar treatment with rituximab and bendamustine,

one patient with a baseline SUVmax of 5.2 died without achieving complete metabolic remission, reflecting no meaningful correlation between SUVmax and survival in this limited number of patients.

Currently, there is no consensus on the optimal discriminatory baseline SUVmax for indolent lymphomas. Schoder et al [26] found that all patients with indolent non-HL had an SUVmax ≤ 13 and that SUVmax >10 was the best cutoff value to discriminate between aggressive and indolent lymphoma with 81% specificity. In our study, we included only confirmed indolent lymphoma cases, and we observed a median SUVmax of 6.35 (2.1 to 15.4) at presentation.

In a prospective study of 38 patients with clinical or histological suspicion of transformed lymphoma, a biopsy of the site with the highest SUVmax showed that all patients with SUVmax >17 had transformed, whereas those with SUVmax <17 had not. The best cutoff value was SUVmax of 14 [27]. In our study, 3 patients underwent transformation during the study period. The SUVmax at transformation ranged from 11 to 19.8 (mean=13.7), which led to a decision of guided biopsy. Although our study size is small, the findings in our study corroborate observations by Barrington et al [28], that FDG-PET/CT can guide targeted biopsies in patients with suspected transformation.

In 2014, the imaging group of the International Conference on Malignant Lymphoma recommended that evaluation of response to treatment of FDG-avid lymphoma should be performed with FDG-PET/CT [28], and FL is considered a lymphoma subtype for which FDG-PET/CT is indicated as mandatory for staging and response assessment. The DC proposed in 2009 are now commonly accepted for PET/CT reporting for response assessment. An advantage of the DC is that a graded response is defined that allows the threshold to be changed according to the proposed intervention: escalation (PET/CT-positive if score ≥ 4) or de-escalation (PET/CT-positive if score ≤ 3) to select true-positive or true-negative patients. It has been proposed that the DC described initially for reporting interim PET/CT may also be used for reporting end-treatment PET/CT [28,29].

The FDG-PET/CT reporting using DC has been used to tailor therapy at the end of treatment and during treatment with a PET/CT-guided strategy in HL and other lymphomas including FL [10,30]. The International Conference on Malignant Lymphoma imaging committee proposed that the DC be extended to end-treatment evaluation [28,31]. With rapid improvements in scanner technology, the minimum score requirement for PET/CT used in the International Harmonization Project (IHP) has become less relevant, and the mediastinal threshold used in the IHP is analogous to a DC score of 1 or 2. Furthermore, a single method of visual assessment in PET/CT is desirable for the interim and end-treatment response assessments. DC scores 1-3 are therefore considered to represent a complete metabolic response under standard treatment. The validity of using the DC with this cutoff for end-treatment PET/CT reporting has recently been confirmed in important studies in FL with a high tumor burden [17].

Federico et al [10] compared the diagnostic accuracy of FDG-PET/CT for clinical evaluation at the end of treatment using the DC and IHP criteria; their results indicated that the

DC is simpler to apply and certainly more effective than the IHP criteria [10].

The results of all the studies discussed above suggest that this holistic approach could improve risk stratification in FL and other indolent lymphomas, and may help build new treatment strategies. However, larger series and prospective multicenter studies are needed before they can be used as risk factors. PET/CT has recently been reported to be useful in detecting BM involvement in FL [32]. However, in our study, of the 17 patients with FL, 12 (70.6%) had BM involvement in the biopsy report, but we could not detect BM involvement in one of those patients, as depicted in Figure 6.

Although using PFS versus overall survival and other quality measures [33] is a matter of debate, we estimated both overall survival and PFS in this prospective study. It seems clear that the growing use of PFS as a primary endpoint in many randomized controlled trials of advanced solid tumors is because a definition for progression exists, we can measure it, and it shortens trial periods. We also used a more meaningful overall survival, but obviously, it requires longer follow-up for a better picture.

The PRIMA trial [34] showed a positive PET/CT scan after treatment, predicted a shorter PFS, and a higher risk of death [24]. A subsequent prospective observational study with a standardized PET/CT acquisition and interpretation criteria confirmed these findings, with a similar risk of progression for PET/CT-positive patients who were assessed centrally by applying a cutoff of ≥ 4 on DC [35].

To compare the central PET/CT review with that of the local investigators for a subset of patients with FL in the PRIMA trial using 2 standardized response criteria, the 2007 IHP criteria and the DC, Tychyj-Pinel, et al [9] reported that at diagnosis, the mean SUVmax of PET-positive cases was 11.7 (range 4.6–35.6), as compared with a mean SUVmax of 8 (2.1–24) in our FL cases. In both studies, there was no significant association between baseline SUVmax and PFS. In 60 postimmunochemotherapy induction scans, Tychyj-Pinel, et al [9], applying the DC with a cutoff ≥ 4 , reported a significantly inferior 42-month PFS in PET-positive patients of 25% versus 61.4% in PET-negative patients ($P=.01$). In our study, only 15 patients with FL underwent postinduction PET/CT scans; 2 patients had a DC score of 5 and their mean PFS was 760 days, whereas 13 patients with a DC score ≤ 3 had a mean PFS of 782.35 days. We observed no significant difference in PFS across these DC scores ($P=.39$). Although in our study population, we could show some trends in overall survival and PFS in relation to SUVmax and DC, we could not obtain the optimal pretreatment SUVmax cutoff for receiver operating characteristic analysis.

Ideas for Future Research

What else can be done from such a study-related data in the future in view of ongoing alternative efforts in the field—food for thought? As we planned our study and started it, multiple

developments in the alternate analysis of PET/CT images have taken place. Although these developments are attractive, due to logistic difficulties, this analysis could not be performed. Some of these ideas are described below in future studies.

Texture analysis and data mining of medical images, also known as *Radiomics*, is an emerging field in radiology. Radiomics is the extraction of occult information from clinical images and subsequent data mining to gain actionable insights. The radiologic texture is the variation in image intensities and occult patterns within an image and is an important part of Radiomics [36,37].

Other exciting developments include integrative FDG-PET/CT, combining a PET/CT scan performed at different time points during treatment or quantitative PET/CT scan parameters with clinical or biological data or other imaging techniques. The integration of information from several different sources (clinical, biological, imaging, etc, along with or integrating with Δ SUVmax [change in SUVmax between 2 points] could lead to more personalized risk stratification) [29].

The prognostic value of total metabolic tumor volume (TMTV) has also been investigated. Trotman et al [17] observed that in patients with FL with a high tumor burden, a TMTV $> 938 \text{ cm}^3$ can identify a small subset of patients with a poor outcome [17].

Total lesion glycolysis, which is the sum of the product of the metabolic volume of each local tumor multiplied by its mean standardized uptake value (total lesion glycolysis = Σ metabolic tumor volume \times mean standardized uptake value), is another parameter that has been studied in DLBCL as a better prognostic tool compared with TMTV [13,14]. It may be attractive to study TMTV in indolent lymphomas.

Finally, keeping the above advancements in mind, our study calls for action in our region for future prospective studies.

Limitations

The limitations of our study include the small sample size divided among the 3 management strategies and a limited period of follow-up in these indolent lymphomas. Making any firm conclusions about the utility of PET/CT based on a small study population and heterogeneous patients is difficult.

Conclusions

In this prospective cohort study consisting of 40 patients, we demonstrated the utility of PET/CT in indolent lymphomas. Overall, the values of SUVmax (< 4.35 vs ≥ 4.35) on interim PET/CT performed the best discriminatory function in predicting PFS. The cumulative PFS of the patients with SUVmax < 4.35 at interim PET/CT, was 100% and the median survival time was 106.67 months, whereas the cumulative PFS of patients with SUVmax ≥ 4.35 was 43.8% and the median survival time was 50.17 months ($P=.04$). The same values showed a good trend in predicting overall survival at baseline and interim PET. Further prospective studies in this region is needed.

Acknowledgments

We would like to thank and acknowledge King Fahad Medical City and their research office for their tremendous support and the Institutional Review Board of King Fahad Medical City for the ethical approval of the study. The authors would also like to thank Mr Mohammed Salman Bashir, Biostatistician Specialist I, Department of Biostatistics, Research Services Administration–Research Center.

Conflicts of Interest

None declared.

References

1. Campo E, Swerdlow SH, Harris NL, Pileri S, Stein H, Jaffe ES. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood* 2011 May 12;117(19):5019-5032 [FREE Full text] [doi: [10.1182/blood-2011-01-293050](https://doi.org/10.1182/blood-2011-01-293050)] [Medline: [21300984](https://pubmed.ncbi.nlm.nih.gov/21300984/)]
2. Solal-Céligny P, Roy P, Colombat P, White J, Armitage JO, Arranz-Saez R, et al. Follicular lymphoma international prognostic index. *Blood* 2004 Sep 01;104(5):1258-1265 [FREE Full text] [doi: [10.1182/blood-2003-12-4434](https://doi.org/10.1182/blood-2003-12-4434)] [Medline: [15126323](https://pubmed.ncbi.nlm.nih.gov/15126323/)]
3. Federico M, Bellei M, Marcheselli L, Luminari S, Lopez-Guillermo A, Vitolo U, et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. *J Clin Oncol* 2009 Sep 20;27(27):4555-4562. [doi: [10.1200/JCO.2008.21.3991](https://doi.org/10.1200/JCO.2008.21.3991)] [Medline: [19652063](https://pubmed.ncbi.nlm.nih.gov/19652063/)]
4. Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* 2002 Jan 24;346(4):235-242. [doi: [10.1056/NEJMoa011795](https://doi.org/10.1056/NEJMoa011795)] [Medline: [11807147](https://pubmed.ncbi.nlm.nih.gov/11807147/)]
5. Morel P, Duhamel A, Gobbi P, Dimopoulos MA, Dhodapkar MV, McCoy J, et al. International prognostic scoring system for Waldenstrom macroglobulinemia. *Blood* 2009 Apr 30;113(18):4163-4170 [FREE Full text] [doi: [10.1182/blood-2008-08-174961](https://doi.org/10.1182/blood-2008-08-174961)] [Medline: [19196866](https://pubmed.ncbi.nlm.nih.gov/19196866/)]
6. Iannitto E, Luminari S, Tripodo C, Mancuso S, Cesaretti M, Marcheselli L, et al. Rituximab with cyclophosphamide, vincristine, non-pegylated liposomal doxorubicin and prednisone as first-line treatment for splenic marginal zone lymphoma: a Fondazione Italiana Linfomi phase II study. *Leuk Lymphoma* 2015 Jul 18;56(12):3281-3287. [doi: [10.3109/10428194.2015.1029925](https://doi.org/10.3109/10428194.2015.1029925)] [Medline: [25791121](https://pubmed.ncbi.nlm.nih.gov/25791121/)]
7. Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the first international workshop on interim-PET-scan in lymphoma. *Leuk Lymphoma* 2009 Aug;50(8):1257-1260. [doi: [10.1080/10428190903040048](https://doi.org/10.1080/10428190903040048)] [Medline: [19544140](https://pubmed.ncbi.nlm.nih.gov/19544140/)]
8. Meignan M, Itti E, Gallamini A, Younes A. FDG PET/CT imaging as a biomarker in lymphoma. *Eur J Nucl Med Mol Imaging* 2015 Apr;42(4):623-633. [doi: [10.1007/s00259-014-2973-6](https://doi.org/10.1007/s00259-014-2973-6)] [Medline: [25573631](https://pubmed.ncbi.nlm.nih.gov/25573631/)]
9. Tychj-Pinel C, Ricard F, Fulham M, Fournier M, Meignan M, Lamy T, et al. PET/CT assessment in follicular lymphoma using standardized criteria: central review in the PRIMA study. *Eur J Nucl Med Mol Imaging* 2014 Mar;41(3):408-415. [doi: [10.1007/s00259-013-2441-8](https://doi.org/10.1007/s00259-013-2441-8)] [Medline: [24435770](https://pubmed.ncbi.nlm.nih.gov/24435770/)]
10. Fallanca F, Alongi P, Incerti E, Gianolli L, Picchio M, Kayani I, et al. Diagnostic accuracy of FDG PET/CT for clinical evaluation at the end of treatment of HL and NHL: a comparison of the Deauville Criteria (DC) and the International Harmonization Project Criteria (IHPC). *Eur J Nucl Med Mol Imaging* 2016 Sep;43(10):1837-1848. [doi: [10.1007/s00259-016-3390-9](https://doi.org/10.1007/s00259-016-3390-9)] [Medline: [27154522](https://pubmed.ncbi.nlm.nih.gov/27154522/)]
11. Gribben J. How I treat indolent lymphoma. *Blood* 2007 Jun 01;109(11):4617-4626 [FREE Full text] [doi: [10.1182/blood-2006-10-041863](https://doi.org/10.1182/blood-2006-10-041863)] [Medline: [17311989](https://pubmed.ncbi.nlm.nih.gov/17311989/)]
12. Lin C, Itti E, Haioun C, Petegnief Y, Luciani A, Dupuis J, et al. Early 18F-FDG PET for prediction of prognosis in patients with diffuse large B-cell lymphoma: SUV-based assessment versus visual analysis. *J Nucl Med* 2007 Oct;48(10):1626-1632 [FREE Full text] [doi: [10.2967/jnumed.107.042093](https://doi.org/10.2967/jnumed.107.042093)] [Medline: [17873129](https://pubmed.ncbi.nlm.nih.gov/17873129/)]
13. Kim TM, Paeng JC, Chun IK, Keam B, Jeon YK, Lee S, et al. Total lesion glycolysis in positron emission tomography is a better predictor of outcome than the International Prognostic Index for patients with diffuse large B cell lymphoma. *Cancer* 2013 Mar 15;119(6):1195-1202 [FREE Full text] [doi: [10.1002/cncr.27855](https://doi.org/10.1002/cncr.27855)] [Medline: [23212736](https://pubmed.ncbi.nlm.nih.gov/23212736/)]
14. Esfahani S, Heidari P, Halpern EF, Hochberg EP, Palmer EL, Mahmood U. Baseline total lesion glycolysis measured with (18)F-FDG PET/CT as a predictor of progression-free survival in diffuse large B-cell lymphoma: a pilot study. *Am J Nucl Med Mol Imaging* 2013 Apr 9;3(3):272-281 [FREE Full text] [Medline: [23638338](https://pubmed.ncbi.nlm.nih.gov/23638338/)]
15. Hutchings M, Barrington SF. PET/CT for therapy response assessment in lymphoma. *J Nucl Med* 2009 May;50 Suppl 1:21-30 [FREE Full text] [doi: [10.2967/jnumed.108.057190](https://doi.org/10.2967/jnumed.108.057190)] [Medline: [19380407](https://pubmed.ncbi.nlm.nih.gov/19380407/)]
16. Bodet-Milin C, Eugène T, Gastinne T, Frampas E, Le Gouill S, Kraeber-Bodéré F. FDG-PET in follicular lymphoma management. *J Oncol* 2012;2012:370272 [FREE Full text] [doi: [10.1155/2012/370272](https://doi.org/10.1155/2012/370272)] [Medline: [22899920](https://pubmed.ncbi.nlm.nih.gov/22899920/)]

17. Trotman J, Luminari S, Boussetta S, Versari A, Dupuis J, Tychyj C, et al. Prognostic value of PET-CT after first-line therapy in patients with follicular lymphoma: a pooled analysis of central scan review in three multicentre studies. *Lancet Haematol* 2014 Oct;1(1):17-27. [doi: [10.1016/S2352-3026\(14\)70008-0](https://doi.org/10.1016/S2352-3026(14)70008-0)] [Medline: [27030064](https://pubmed.ncbi.nlm.nih.gov/27030064/)]
18. Lu Z, Lin M, Downe P, Chong S, Ling S. The prognostic value of mid- and post-treatment [(18)F] fluorodeoxyglucose (FDG) positron emission tomography (PET) in indolent follicular lymphoma. *Ann Nucl Med* 2014 Oct;28(8):805-811. [doi: [10.1007/s12149-014-0874-1](https://doi.org/10.1007/s12149-014-0874-1)] [Medline: [25008291](https://pubmed.ncbi.nlm.nih.gov/25008291/)]
19. El-Galaly TC, Gormsen LC, Hutchings M. PET/CT for staging; past, present, and future. *Semin Nucl Med* 2018 Jan;48(1):4-16. [doi: [10.1053/j.semnuclmed.2017.09.001](https://doi.org/10.1053/j.semnuclmed.2017.09.001)] [Medline: [29195617](https://pubmed.ncbi.nlm.nih.gov/29195617/)]
20. Sarkozy C, Salles G, Bachy E. Chemotherapy-free treatment in patients with follicular lymphoma. *Expert Rev Hematol* 2015 Apr;8(2):187-203. [doi: [10.1586/17474086.2015.998995](https://doi.org/10.1586/17474086.2015.998995)] [Medline: [25585961](https://pubmed.ncbi.nlm.nih.gov/25585961/)]
21. Karmali R, Kimby E, Ghielmini M, Flinn I, Gordon L, Zucca E. Rituximab: a benchmark in the development of chemotherapy-free treatment strategies for follicular lymphomas. *Ann Oncol* 2018 Feb 01;29(2):332-340 [FREE Full text] [doi: [10.1093/annonc/mdx768](https://doi.org/10.1093/annonc/mdx768)] [Medline: [29206891](https://pubmed.ncbi.nlm.nih.gov/29206891/)]
22. Casulo C. How I manage patients with follicular lymphoma. *Br J Haematol* 2019 Aug;186(4):513-523. [doi: [10.1111/bjh.16011](https://doi.org/10.1111/bjh.16011)] [Medline: [31173345](https://pubmed.ncbi.nlm.nih.gov/31173345/)]
23. Lockmer S, Østenstad B, Hagberg H, Holte H, Johansson A, Wahlin BE, et al. Chemotherapy-free initial treatment of advanced indolent lymphoma has durable effect with low toxicity: results from two nordic lymphoma group trials with more than 10 years of follow-up. *J Clin Oncol* 2018 Oct 04:00262 (forthcoming). [doi: [10.1200/JCO.18.00262](https://doi.org/10.1200/JCO.18.00262)] [Medline: [30285560](https://pubmed.ncbi.nlm.nih.gov/30285560/)]
24. Trotman J, Fournier M, Lamy T, Seymour JF, Sonet A, Janikova A, et al. Positron emission tomography-computed tomography (PET-CT) after induction therapy is highly predictive of patient outcome in follicular lymphoma: analysis of PET-CT in a subset of PRIMA trial participants. *J Clin Oncol* 2011 Aug 10;29(23):3194-3200. [doi: [10.1200/JCO.2011.35.0736](https://doi.org/10.1200/JCO.2011.35.0736)] [Medline: [21747087](https://pubmed.ncbi.nlm.nih.gov/21747087/)]
25. Karam M, Ata A, Irish K, Feustel PJ, Mottaghy FM, Stroobants SG, et al. FDG positron emission tomography/computed tomography scan may identify mantle cell lymphoma patients with unusually favorable outcome. *Nucl Med Commun* 2009 Oct;30(10):770-778. [doi: [10.1097/MNM.0b013e32832e0c13](https://doi.org/10.1097/MNM.0b013e32832e0c13)] [Medline: [19657307](https://pubmed.ncbi.nlm.nih.gov/19657307/)]
26. Schöder H, Noy A, Gönen M, Weng L, Green D, Erdi YE, et al. Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma. *J Clin Oncol* 2005 Jul 20;23(21):4643-4651. [doi: [10.1200/JCO.2005.12.072](https://doi.org/10.1200/JCO.2005.12.072)] [Medline: [15837966](https://pubmed.ncbi.nlm.nih.gov/15837966/)]
27. Bodet-Milin C, Kraeber-Bodéré F, Moreau P, Campion L, Dupas B, Le Gouill S. Investigation of FDG-PET/CT imaging to guide biopsies in the detection of histological transformation of indolent lymphoma. *Haematologica* 2008 Mar;93(3):471-472 [FREE Full text] [doi: [10.3324/haematol.12013](https://doi.org/10.3324/haematol.12013)] [Medline: [18310543](https://pubmed.ncbi.nlm.nih.gov/18310543/)]
28. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Müeller SP, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol* 2014 Sep 20;32(27):3048-3058 [FREE Full text] [doi: [10.1200/JCO.2013.53.5229](https://doi.org/10.1200/JCO.2013.53.5229)] [Medline: [25113771](https://pubmed.ncbi.nlm.nih.gov/25113771/)]
29. Meignan M, Gallamini A, Itti E, Barrington S, Haioun C, Polliack A. Report on the Third International Workshop on interim positron emission tomography in lymphoma held in Menton, France, 26-27 September 2011 and Menton 2011 consensus. *Leuk Lymphoma* 2012 Oct;53(10):1876-1881. [doi: [10.3109/10428194.2012.677535](https://doi.org/10.3109/10428194.2012.677535)] [Medline: [22432519](https://pubmed.ncbi.nlm.nih.gov/22432519/)]
30. Hasenclever D, Kurch L, Mauz-Körholz C, Elsner A, Georgi T, Wallace H, et al. qPET - a quantitative extension of the Deauville scale to assess response in interim FDG-PET scans in lymphoma. *Eur J Nucl Med Mol Imaging* 2014 Jul;41(7):1301-1308. [doi: [10.1007/s00259-014-2715-9](https://doi.org/10.1007/s00259-014-2715-9)] [Medline: [24604592](https://pubmed.ncbi.nlm.nih.gov/24604592/)]
31. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, Alliance, Australasian Leukaemia Lymphoma Group, Eastern Cooperative Oncology Group, European Mantle Cell Lymphoma Consortium, Italian Lymphoma Foundation, European Organisation for Research, Treatment of Cancer/Dutch Hemato-Oncology Group, Grupo Español de Médula Ósea, German High-Grade Lymphoma Study Group, German Hodgkin's Study Group, Japanese Lymphoma Study Group, Lymphoma Study Association, NCIC Clinical Trials Group, Nordic Lymphoma Study Group, Southwest Oncology Group, United Kingdom National Cancer Research Institute. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol* 2014 Sep 20;32(27):3059-3068 [FREE Full text] [doi: [10.1200/JCO.2013.54.8800](https://doi.org/10.1200/JCO.2013.54.8800)] [Medline: [25113753](https://pubmed.ncbi.nlm.nih.gov/25113753/)]
32. Nakajima R, Moskowitz AJ, Michaud L, Mauguen A, Batlevi CL, Dogan A, et al. Baseline FDG-PET/CT detects bone marrow involvement in follicular lymphoma and provides relevant prognostic information. *Blood Adv* 2020 Apr 28;4(8):1812-1823 [FREE Full text] [doi: [10.1182/bloodadvances.2020001579](https://doi.org/10.1182/bloodadvances.2020001579)] [Medline: [32343798](https://pubmed.ncbi.nlm.nih.gov/32343798/)]
33. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol* 2012 Apr 01;30(10):1030-1033. [doi: [10.1200/JCO.2011.38.7571](https://doi.org/10.1200/JCO.2011.38.7571)] [Medline: [22370321](https://pubmed.ncbi.nlm.nih.gov/22370321/)]
34. Salles G, Seymour JF, Offner F, López-Guillermo A, Belada D, Xerri L, et al. Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial. *Lancet* 2011 Jan 01;377(9759):42-51. [doi: [10.1016/S0140-6736\(10\)62175-7](https://doi.org/10.1016/S0140-6736(10)62175-7)] [Medline: [21176949](https://pubmed.ncbi.nlm.nih.gov/21176949/)]

35. Dupuis J, Berriolo-Riedinger A, Julian A, Brice P, Tychyj-Pinel C, Tilly H, et al. Impact of [(18)F]fluorodeoxyglucose positron emission tomography response evaluation in patients with high-tumor burden follicular lymphoma treated with immunochemotherapy: a prospective study from the Groupe d'Etudes des Lymphomes de l'Adulte and GOELAMS. *J Clin Oncol* 2012 Dec 10;30(35):4317-4322. [doi: [10.1200/JCO.2012.43.0934](https://doi.org/10.1200/JCO.2012.43.0934)] [Medline: [23109699](https://pubmed.ncbi.nlm.nih.gov/23109699/)]
36. Varghese BA, Cen SY, Hwang DH, Duddalwar VA. Texture analysis of imaging: what radiologists need to know. *AJR Am J Roentgenol* 2019 Mar;212(3):520-528. [doi: [10.2214/AJR.18.20624](https://doi.org/10.2214/AJR.18.20624)] [Medline: [30645163](https://pubmed.ncbi.nlm.nih.gov/30645163/)]
37. Lambin P, Zindler J, Vanneste BG, De Voorde LV, Eekers D, Compter I, et al. Decision support systems for personalized and participative radiation oncology. *Adv Drug Deliv Rev* 2017 Jan 15;109:131-153. [doi: [10.1016/j.addr.2016.01.006](https://doi.org/10.1016/j.addr.2016.01.006)] [Medline: [26774327](https://pubmed.ncbi.nlm.nih.gov/26774327/)]

Abbreviations

- BM:** bone marrow
CLL: chronic lymphocytic leukemia
CT: computed tomography
DC: Deauville criteria
DLBCL: diffuse large B-cell lymphoma
FDG: fluorodeoxyglucose
FL: follicular lymphoma
FLIPI: follicular lymphoma international prognostic index
HL: Hodgkin lymphoma
IHP: International Harmonization Project
KM: Kaplan–Meier
MCL: mantle cell lymphoma
MIP: maximum intensity projection
PET: positron emission tomography
PFS: progression-free survival
R-CHOP: rituximab-cyclophosphamide, hydroxydaunorubicin, vincristine and prednisone
R-CVP: rituximab-cyclophosphamide, vincristine and prednisone
RFS: relapse-free survival
SLL: small lymphocytic lymphoma
SUVmax: maximum standardized uptake value
TMTV: total metabolic tumor volume

Edited by G Eysenbach; submitted 10.10.20; peer-reviewed by D Brotman, A Signore, E De Bernardi; comments to author 11.05.21; revised version received 23.07.21; accepted 07.09.21; published 12.11.21.

Please cite as:

AlShehry NF, Shanker R, Zaidi SZA, AlGhmlas F, Motabi IH, Iqbal S, Butt AA, AlShehri H, Tailor IK, Altaf SY, AlGhamdi M, Marie M, AlFayez M, Al Zahrani K, Dwaimah M, Al-Halouli T, Al-Shakweer W, AlShehery MZ, Zaidi ARZ, Gill AM, Albutosh BM, Ahmed M

Role of 18F-Fluorodeoxyglucose–Positron Emission Tomography/Computed Tomography Imaging in the Prediction of Prognosis in Patients With Indolent Lymphoma: Prospective Study

JMIR Form Res 2021;5(11):e24936

URL: <https://formative.jmir.org/2021/11/e24936>

doi: [10.2196/24936](https://doi.org/10.2196/24936)

PMID: [34508363](https://pubmed.ncbi.nlm.nih.gov/34508363/)

©Nawal Faiez AlShehry, Raja Shanker, Syed Ziauddin Ahmed Zaidi, Fahad AlGhmlas, Ibraheem Hussein Motabi, Shahid Iqbal, Ahmad Ali Butt, Hassan AlShehri, Imran Khan Tailor, Syed Yasir Altaf, Mubarak AlGhamdi, Mohammed Marie, Mansour AlFayez, Kamal Al Zahrani, Mohammed Dwaimah, Tahani Al-Halouli, Wafaa Al-Shakweer, Maied Zaher AlShehery, Abdul Rehman Zia Zaidi, Atta Munawar Gill, Belal Mohammed Albutosh, Musab Ahmed. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.11.2021. 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

Implementation Outcome Scales for Digital Mental Health (iOSDMH): Scale Development and Cross-sectional Study

Natsu Sasaki^{1*}, MD; Erika Obikane^{1*}, MD, MPH, PhD; Rajesh Vedanthan², MD, PhD; Kotaro Imamura¹, MPH, PhD; Pim Cuijpers³, PhD; Taichi Shimazu⁴, MD, PhD; Masamitsu Kamada⁵, PhD; Norito Kawakami¹, MD, PhD; Daisuke Nishi¹, MD, PhD

¹Department of Mental Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

²Department of Population Health, Grossman School of Medicine, New York University, New York, NY, United States

³Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

⁴Behavioral Science Division, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan

⁵Department of Health and Social Behavior, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

*these authors contributed equally

Corresponding Author:

Daisuke Nishi, MD, PhD

Department of Mental Health

Graduate School of Medicine

The University of Tokyo

7-3-1 Hongo, Bunkyo-ku

Tokyo, 113-0033

Japan

Phone: 81 3 5841 3521

Email: d-nishi@m.u-tokyo.ac.jp

Abstract

Background: Digital mental health interventions are being used more than ever for the prevention and treatment of psychological problems. Optimizing the implementation aspects of digital mental health is essential to deliver the program to populations in need, but there is a lack of validated implementation outcome measures for digital mental health interventions.

Objective: The primary aim of this study is to develop implementation outcome scales of digital mental health for different levels of stakeholders involved in the implementation process: users, providers, and managers or policy makers. The secondary aim is to validate the developed scale for users.

Methods: We developed English and Japanese versions of the implementation outcome scales for digital mental health (iOSDMH) based on the literature review and panel discussions with experts in implementation research and web-based psychotherapy. The study developed acceptability, appropriateness, feasibility, satisfaction, and harm as the outcome measures for users, providers, and managers or policy makers. We conducted evidence-based interventions via the internet using UTSMed, a website for mental health information (N=200). Exploratory factor analysis (EFA) was conducted to assess the structural validity of the iOSDMH for users. Satisfaction, which consisted of a single item, was not included in the EFA.

Results: The iOSDMH was developed for users, providers, and managers or policy makers. The iOSDMH contains 19 items for users, 11 items for providers, and 14 items for managers or policy makers. Cronbach α coefficients indicated intermediate internal consistency for acceptability ($\alpha=.665$) but high consistency for appropriateness ($\alpha=.776$), feasibility ($\alpha=.832$), and harm ($\alpha=.777$) of the iOSDMH for users. EFA revealed 3-factor structures, indicating acceptability and appropriateness as close concepts. Despite the similarity between these 2 concepts, we inferred that acceptability and appropriateness should be used as different factors, following previous studies.

Conclusions: We developed iOSDMH for users, providers, and managers. Psychometric assessment of the scales for users demonstrated acceptable reliability and validity. Evaluating the components of digital mental health implementation is a major step forward in implementation science.

(JMIR Form Res 2021;5(11):e24332) doi:[10.2196/24332](https://doi.org/10.2196/24332)

KEYWORDS

implementation outcomes; acceptability; appropriateness; feasibility; harm

Introduction

Background

Due to rapid advances in technology, mental health interventions delivered using digital and telecommunication technologies have become an alternative to face-to-face interventions. Digital mental health interventions vary from teleconsultation with specialists (eg, physicians, nurses, psychotherapists) to fully or partially automated programs led by web-based systems or artificial intelligence [1,2]. For example, internet-based cognitive behavioral therapy has been found useful for improving depression, anxiety disorders, and other psychiatric conditions [3-5]. Moreover, a recent meta-analysis suggested that internet-based interventions were effective in preventing the onset of depression among individuals with subthreshold depression, indicating future implications for community prevention [6]. Past studies have demonstrated that mental health interventions are suitable for digital platforms because of several reasons: rare need for laboratory testing of patients, chronic shortage of human resources in the field of mental health, and stigma often experienced by patients in consulting mental health professionals [7].

Although numerous studies have demonstrated the efficacy of digital mental health interventions, many people do not benefit from them mainly due to insufficient implementation. Implementation is defined as “a specified set of activities designed to put into practice a policy or intervention of known dimensions” [8]. The entire care cascade can benefit from optimization. People with mental health problems are known to face psychological obstacles to treatment [9] due to lack of motivation [9,10], lower mental literacy [11], or stigma [12]. Moreover, digital mental health interventions face high attrition and low adherence to programs especially in open-access websites [13-15]. This may be because implementation aspects have not been fully examined when the interventions are being developed. One of the major barriers is the lack of reliable and valid process measures. Validated measures are needed to monitor and evaluate implementation efforts. Core implementation outcomes include acceptability, appropriateness, feasibility, adoption, penetration, cost, fidelity, and sustainability [16,17]. However, most of these measures have not yet been validated. Weiner et al have developed validated scales for acceptability, appropriateness, and feasibility [18], but these scales were not designed for digital mental health settings. A systematic review of implementation outcomes in mental health settings reported that most outcomes focused on acceptability, and other constructs were underdeveloped without psychometric assessment [19].

Moreover, implementation involves not only the patients targeted by an intervention but also individuals or groups responsible for program management, including health care providers, policy makers, and community-based organizations [8]. Providers have direct contact with users. Managers or policy

makers have the authority to decide on the implementation of these programs.

Objectives

To our best knowledge, outcome measurements to evaluate implementation aspects concerning users, providers, and managers or policy makers are not available in digital mental health research. Therefore, the primary aim of this study is to develop new implementation outcome scales for digital mental health (iOSDMH) interventions that can be applied for users, providers, and managers or policy makers. The secondary aim is to validate the implementation scale for users. This study does not include validation of the implementation scale for providers and managers because the study does not involve providers and managers.

Methods

Study Design

We originally developed the English and Japanese versions of the iOSDMH based on previously published literature [18,19], which proposed the 3 measures of the implementation outcome scale and provided a systematic review of implementation outcomes. The development of iOSDMH consisted of 3 phases. In the first phase, literature review on implementation scales was conducted, and scales with high scores for evidence-based criteria were selected for further review. Each item from the item pool was critically reviewed by 3 researchers, and they discussed whether the items were relevant for digital mental health. Based on the selected items, the team developed the first drafts of the scales for users, providers, and managers or policy makers. In the second phase, the draft of the iOSDMH was carefully examined by 2 implementation researchers and 1 mental health researcher. With these expert panels, the research team discussed the relevance of the selected items in each category as well as the wording of each question and created the second drafts of the scales. In the third phase, the draft of the iOSDMH was presented to the implementation and digital mental health researchers to confirm the scales and further changes were made based on their inputs. After confirming the relevance of the scales with the expert panels, we conducted an internet-based survey to examine the scale properties of the Japanese version of the iOSDMH for users. Although the iOSDMH targeted 3 categories of implementation stakeholders, namely users, providers, and managers or policy makers [8], tool validation was conducted for users only, as the study did not involve providers and managers.

Ethical Considerations

This study was approved by The Research Ethics Committee of the Graduate School of Medicine/Faculty of Medicine, University of Tokyo (No. 2019361NI). The aims and procedures of the study were explained on the web page before participants answered the questionnaire. Responses to the questionnaire were considered as the consent to participate.

Development Process of iOSDMH

The development of the iOSDMH consisted of 3 phases. In the first phase, 3 of the investigators (EO, NS, and DN) reviewed 89 implementation scales from previous literature and a systematic review of implementation outcomes [18,19]. After the review, we selected 9 implementation scales (171 items) that were rated with evidence-based criteria in the following categories: acceptability of the intervention process, acceptability of the implementation process, adoption, cost, feasibility, penetration, and sustainability. Each item was reviewed carefully by 3 researchers, and 4 highly scored instruments in terms of psychometric and pragmatic quality were selected [20-23]. The following concepts were considered relevant in measuring implementation aspects of digital mental health interventions. Moore et al [21] developed the assessment tool for adoption of technology interventions. Whittingham et al [22] evaluated the acceptability of the parent training program. Hides et al [20] reported the feasibility and acceptability of mental health training for alcohol and drug use. Yetter [23] reported the acceptability of psychotherapeutic intervention in schools. Relevant items were adapted for the web-based mental health interventions, and those not relevant in the context of digital mental health were excluded.

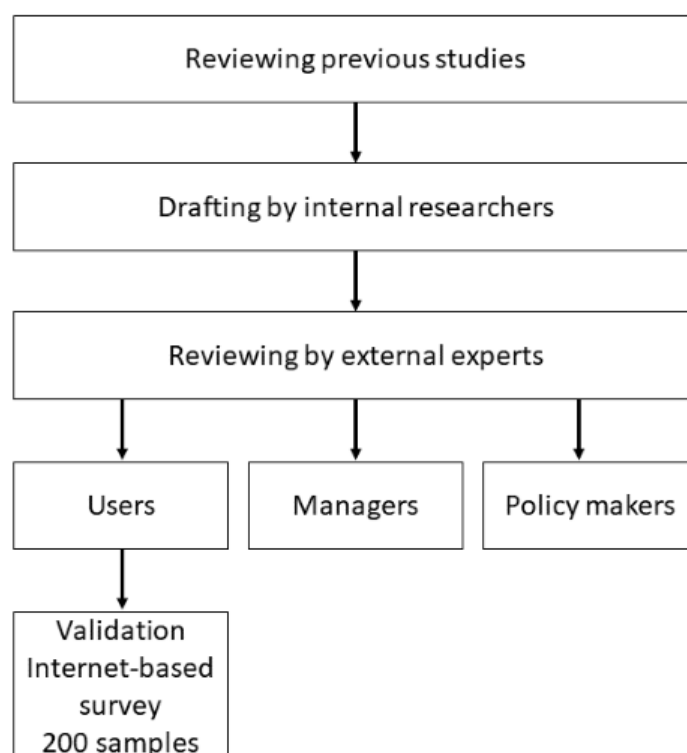
The iOSDMH consisted of two parts: (1) evaluations and (2) adverse events of using digital mental health programs.

In the second phase, the drafts of the iOSDMH for users, providers, and managers were reviewed by experts on web-based psychotherapy (KI) and implementation science (MK and RV), and a consensus was reached to categorize all items into the concepts of acceptability, appropriateness, and feasibility for evaluation. We primarily had 22 items for evaluating the use of digital mental health programs and 6 adverse events of the program for users. We narrowed these to 14 items for evaluations and 5 items for adverse events following discussions with expert panels. For the iOSDMH of providers, we first had 14 items for evaluations and 1 item for adverse events; we then selected 10 items for evaluations and 1 item for adverse events. For the iOSDMH of managers, we first had 11 items for evaluations and 1 item for adverse events but changed them to 13 items for evaluations and 1 item for adverse events. Acceptability is the perception that a given practice is agreeable or palatable, such as feeling "I like it." Wordings of the items on acceptability (Items 1, 2, and 3 for users, and Item 2 for managers) were taken from Moore et al [21]. Item 3 for users and Items 1, 3, and 4 for managers were from Whittingham et al [22]. The wording of Item 4 for providers was from Yetter et al [23]. Appropriateness is the perceived fit, relevance, or compatibility, such as feeling "I think it is right to do." Wordings of Item 5 for users, and Items 5 and 7 for managers were from Moore et al [21]. The wording of Item 8 for providers was based on Hides et al [20]. Items 4, 6, and 7 for users and Item 6 for managers were originally developed based on discussions. Item 9 for providers and Item 8 for managers were worded according to Whittingham et al [22]. Feasibility is the extent to which a

practice can be successfully implemented [17]. Wordings of Items 8 and 9 for users, Item 7 for providers, and Item 12 for managers were from Moore et al [21]. Items 10, 11, and 13 for users and Item 8 for providers were from Hides et al [20]. Items 12 and 14 for users, Item 9 for providers, and Items 9, 10 and 11 for managers were originally developed based on discussions. In addition to the 3 concepts, we added 1 item related to overall satisfaction in the evaluation section because overall satisfaction is considered important in implementation processes [17]. Previous literature distinguished between satisfaction and acceptability, with acceptability being a more specific concept referring to a certain intervention and satisfaction usually representing general experience [16]. However, we considered that overall satisfaction was an important client outcome of process measures. The second part involved harm (ie, adverse effects of interventions). Burdens and adverse events in using digital programs should be considered because digital mental health interventions are not harm free [24].

In the final step, the second drafts of the iOSDMH for users, providers, and managers were reviewed by 2 external researchers (PC and TS), 1 digital mental health researcher, and 1 implementation researcher, and corrections were made based on discussions. We recognized that the relevance of some items differed according to cultural contexts of responders. For example, Item 2 on acceptability for users, Item 3 on acceptability for providers, and Item 2 on acceptability for managers asked whether using the program would improve their social image, or their evaluation of themselves or their organizations. Improving social image may be important and beneficial in some cultural groups but not as much in others. Researchers of 3 different countries considered these items to be relevant, and therefore, we preserved these items. All coauthors engaged in a series of discussions until a consensus was reached on whether the items reflected the appropriate concepts, as well as the overall comprehensiveness and relevance of the scale. None of the objective criteria was adopted in the process of reaching consensus.

The iOSDMH was developed for targeting 3 different groups that are involved in the implementation process: users (ie, patients), providers, and managers or policy makers. Providers are people who have direct contact with users (eg, medical: nurse; workplace: person in charge). Managers or policy makers are people who have authority to decide on the implementation of this program (eg, responsible person). These scales did not restrict the study settings (eg, clinic workplace, and school). For example, the implementation of workplace-based interventions may involve workers (users), human resource staff (providers), and company owners (policy makers). Moreover, these scales aimed to evaluate the implementation aspects related to users, providers, and managers after the users completed or at least partially received the internet-based intervention. Most items were developed assuming that the users had prior experience in receiving the internet-based intervention. The process of developing the iOSDMH is shown in Figure 1.

Figure 1. Development process of the implementation outcome scales for digital mental health.

Internet-Based Survey

Participants were recruited on an internet-based crowd working system (CrowdWorks, Inc), which has more than 2 million registered workers. The criterion for eligibility was to be over 20 years old. Participants were required to learn from the self-help information website UTSMed [25], a digital mental health intervention. The UTSMed website was developed to help Japanese general workers cope with stress and depression. It contains self-learning psychoeducational information on mental health (eg, stress management). This web-based UTSMed intervention has proven effective in reducing depressive symptoms and improving work engagement among Japanese workers in previous randomized controlled trials [26,27]. In our study, participants were asked to explore the UTSMed website for as long as they liked and take quizzes on mental health. They answered the Japanese version of the iOSDMH for users (14 items in 2 pages) after they received acceptable scores (ie, 8 or more of 10 questions answered correctly) in the quizzes. The participants received web-based points as incentives for participation. As the current UTSMed is an open-access website and authors directly provided the URL to participants, the study did not involve any providers, managers, or policy makers. The psychometric assessment thus was limited to users. Gender, age, marital status, education attainment, income, work status, occupation type, and employment contract constituted the demographic information. The target sample size was determined as 10 times the number of items needed to obtain reliable results (eg, 200 participants). The survey was conducted through the internet-based crowd working system. Completed answers were obtained without missing variables.

Statistical Analysis

To assess the internal consistency of the Japanese iOSDMH, Cronbach α coefficients were calculated for all scales and each of the 4 subscales (acceptability, appropriateness, feasibility, and harm). To assess structural validity, exploratory factor analysis (EFA) was conducted because previous studies have shown that acceptability and appropriateness are conceptually similar [16,18]. EFA was conducted by excluding 1 item of overall satisfaction, as the concept of satisfaction cannot be applied to each of the 4 subscales. We extracted factors with eigenvalues of more than 1, following the Kaiser–Guttman “eigenvalues greater than one” criterion [28], using the least-squares method with Promax rotation. Items with factor loadings above 0.4 were retained [29].

Statistical significance was defined as $P < .05$. All statistical analyses were performed using the Japanese version of SPSS 26.0 (IBM Corp).

Results

Development of iOSDMH

The final version of the iOSDMH for users contained 3 items for acceptability based on Moore and Whittingham [21,22] and 4 items for appropriateness, 1 of which was based on Moore [21]. The others were original; there were 6 items for feasibility, 5 of which were based on Moore and Hide [20,21], and 1 item was original; we developed 5 original items for harm and 1 for overall satisfaction. The iOSDMH states “Please read the following statements and select ONE option that most describes your opinion about the program.” The response to each item was scored on a 4-point Likert-type scale ranging from 1 (disagree) to 4 (agree). The iOSDMH for providers and

managers or policy makers has an option 5 (don't know). Details are provided in [Multimedia Appendix 1](#).

The final version of the iOSDMH for providers contained 3 items for acceptability, 2 of which were based on Yetter [23], and 1 item was original; 3 items for appropriateness, 2 of which were based on Yetter [23], and 1 item was original; 3 items for feasibility, 1 of which was original and 2 were based on Moore [21] and Hides [20]; 1 original item for harm; and 1 for overall satisfaction. For acceptability, Item 1 evaluated the providers' perceived acceptance of the program for protecting the mental health of its users, whereas Items 2 and 3 focused on their own acceptability to implement the program in their workplace. For appropriateness, Items 4 and 6 asked about the providers' perceived appropriateness of the program for users, whereas Item 5 asked about the appropriateness of the program considering the situation of the providers. For feasibility, Item 7 evaluated the providers' perception of the program's feasibility for users, and Items 8 and 9 focused on the willingness of providers to provide the program to users.

The final version of the iOSDMH for managers or policy makers contained 4 items for acceptability, 3 of which were based on Whittingham [22] and the other on Moore [21]; 4 items were for appropriateness, 2 of which were based on Moore [21]; 1 item was based on Whittingham [22], and the other one was

original; there were 4 items for feasibility, 1 of which was based on Moore [21] and the others were original; we had 1 original item for harm and 1 item for overall satisfaction. Similar to the iOSDMH for providers, each factor of the scale contained questions on managers' perceptions on implementation in terms of the conditions of users and providers, as well as the managers themselves. For example, Items 1 and 2 asked about the acceptability of the program for the institution, whereas Item 3 focused on managers' perceived acceptability for providers, and Item 4 evaluated managers' perceived acceptability for users. For appropriateness, Items 5 and 7 focused on the appropriateness of the program for the institution, and Item 8 assessed the appropriateness of the program for users according to managers' perceptions. For feasibility, Items 9 and 10 examined the feasibility of the program for the institution as perceived by managers or policy makers. Item 11 evaluated managers' perceived feasibility for providers, and Item 12 evaluated managers' perception of feasibility for users.

Internet-Based Survey

We recruited 200 participants, whose characteristics are presented in [Table 1](#). Most were female (n=110, 55%), single (n=100, 50%), had an undergraduate education (n=114, 57%), and were employed (n=156, 78%). Their average age was 39.18 years (SD 9.81), with the minimum age being 20 years and the maximum being 76 years.

Table 1. Participant characteristics obtained from the internet-based survey (N=200).

Participant characteristics	n (%)
Gender	
Male	90 (45)
Female	110 (55)
Age, years	
20 to 29	32 (16)
30 to 39	78 (39)
40 to 49	59 (29)
Over 50	30 (15)
Not mentioned	1 (0.5)
Marital status	
Single	99 (49.5)
Married	93 (46.5)
Divorced/widowed	8 (4)
Education	
Junior high school	2 (1)
High school	40 (20)
College/vocational school	37 (18.5)
Undergraduate	114 (57)
Postgraduate	7 (3.5)
Individual income	
No income	31 (15.5)
<2 million yen	69 (34.5)
2 to 4 million yen	48 (24)
4 to 6 million yen	34 (17)
6 to 8 million yen	13 (6.5)
8 to 10 million yen	5 (2.5)
Work status	
Employed	155 (77.5)
Unemployed	45 (22.5)
Occupation type	
Managers	8 (4)
Specialists/technicians	31 (15.5)
Office work	37 (18.5)
Manual work	19 (9.5)
Service/marketing	21 (10.5)
Others	42 (21)
Unemployed	42 (21)
Employment contract	
Full-time	69 (34.5)
Contract worker	16 (8)
Temporary worker	6 (3)
Part-time	28 (14)

Participant characteristics	n (%)
Self-employed	32 (16)
Others	6 (3)
Unemployed	43 (21.5)

Internal Consistency

Table 2 shows the mean scores of the iOSDMH for users and Cronbach α values. The mean of the total score of the iOSDMH

was 51.73 (range 19-76). The Cronbach α values were slightly below the threshold ($\alpha > .7$) for acceptability ($\alpha = .665$), but well above the threshold for appropriateness ($\alpha = .776$), feasibility ($\alpha = .832$), and harm ($\alpha = .777$).

Table 2. Average, SD, and reliability among the Japanese population for the iOSDMH and their subscales (N=200).

iOSDMH ^a subscales (number of items; possible range)	Mean (SD)	Cronbach α
Total (19 items; 19-76)	51.73 (5.1)	.685
Acceptability (3 items; 3-12)	8.62 (2.43)	.665
Appropriateness (4 items; 4-16)	11.76 (4.21)	.776
Feasibility (6 items; 6-24)	18.84 (7.94)	.832
Harm (5 items; 5-20)	9.47 (8.64)	.777
Satisfaction (1 item; 1-4)	3.06 (0.58)	N/A ^b

^aiOSDMH: implementation outcome scales for digital mental health.

^bNot available.

Factor Structure of iOSDMH

The EFA results are shown in Table 3. EFA conducted according to the Kaiser–Guttman criterion yielded 3 factors. The first

factors were acceptability and appropriateness. The second was feasibility, and the third was harm. All items showed factor loadings above 0.4, so we kept them intact.

Table 3. Exploratory factor analysis without assuming the number of factors by using least-squares method with Promax rotation^a.

Item number	Short description of the item	Concept	Factor loading score		
			1	2	3
6	Suitable for my social conditions	Appropriateness	<i>0.813</i>	-0.192	-0.063
5	Applicable to my health status	Appropriateness	<i>0.757</i>	-0.076	0.089
7	Fits my living condition	Appropriateness	<i>0.696</i>	-0.090	-0.130
3	Acceptable for me	Acceptability	<i>0.695</i>	0.174	-0.011
2	Improves my social image	Acceptability	<i>0.532</i>	-0.017	-0.037
1	Advantages outweigh the disadvantages for keeping my mental health	Acceptability	<i>0.481</i>	0.019	0.067
4	Appropriate (from your perspective, it is the right thing to do)	Appropriateness	<i>0.463</i>	0.325	0.074
10	Total length is implementable	Feasibility	-0.135	<i>0.922</i>	0.11
11	Length of one content is implementable	Feasibility	-0.113	<i>0.839</i>	-0.039
12	Frequency is implementable	Feasibility	0.073	<i>0.575</i>	-0.039
13	Easy to understand	Feasibility	0.163	<i>0.559</i>	-0.009
9 ^b	Physical effort	Feasibility	-0.062	<i>0.518</i>	-0.24
8	Easy to use	Feasibility	0.372	<i>0.492</i>	0.124
18	Time-consuming	Harm	-0.075	-0.086	<i>0.723</i>
16	Mental symptoms	Harm	-0.009	-0.154	<i>0.657</i>
17	Induced dangerous experience regarding safety	Harm	0.051	0.227	<i>0.610</i>
15	Physical symptoms	Harm	0.086	-0.235	<i>0.592</i>
19	Excessive pressure on learning regularly	Harm	-0.077	-0.019	<i>0.562</i>

^aItalicized values are significant.

^bUsed a reversed score.

Discussion

Principal Findings

This study developed implementation outcome scales for digital mental health based on existing literature and reviews by experts on web-based psychotherapy and implementation science. Our measurements included 3 key constructs of the implementation outcomes (acceptability, appropriateness, and feasibility) from previous studies and additional constructs on harm and satisfaction considered necessary in the implementation process. Implementation researchers and mental health experts agreed that each instrument of the implementation measures reflected the correct concepts.

This study created implementation outcomes for people involved in the implementation process: users, providers, and managers or policy makers. According to the World Health Organization's implementation research guide, knowledge exchange or collaborative problem-solving should occur among stakeholders such as providers, managers or policy makers, and researchers [8]. A past study indicated that policy makers and primary stakeholders had decision frameworks that would produce different implementation outcomes [30]. Previous implementation outcome research targeted 1 or 2 groups of users, providers, and managers or policy makers. However, to our knowledge, few studies have resulted in outcome scales for

different levels of stakeholders [19]. We believed that outcome measures should be adjusted for stakeholders, as decision frameworks may differ among them. For example, users of the program judge its appropriateness by considering whether it is suitable for their situations. Nevertheless, providers may find the program suitable for the circumstances of their users and for themselves. Managers or policy makers will care if the program is suitable for themselves and for users and providers. Another example is that although the length or frequency of the program may be important for feasibility among users, the cost or institutional resources may be important in assessing feasibility among managers or policy makers.

Psychometric assessment of the implementation outcomes showed good internal consistency for appropriateness, feasibility, and harm. Internal consistency for acceptability was lower than that for other constructs ($\alpha=.665$), possibly because the construct for acceptability consisted of only 3 items. The EFA model suggested a 3-factor solution. The first factors were acceptability and appropriateness. Correlations between these 2 concepts were high. Our finding was consistent with previous studies in that acceptability and appropriateness were conceptually close [16,18]. For instance, it has been reported that perceived acceptability of treatment is shaped by factors such as appropriateness, suitability, convenience, and effectiveness [31,32]. However, other scholars agree that acceptability should be distinguished from appropriateness.

Proctor et al stated that an individual (ie, end user) may think that an intervention that seems appropriate may not always be acceptable and vice versa [16]. Similarly, previous research on alcohol screening in emergency departments revealed that nurses and physicians found alcohol screening to be acceptable but not appropriate because the process was time-consuming, the patients might object to it, and the nurses had not received sufficient training [33,34]. Therefore, it is essential to distinguish acceptability from appropriateness in such a situation because it helps focus on the appropriate concept during implementation. Therefore, we decided to maintain the 4-factor questionnaire comprising acceptability, appropriateness, feasibility, and harm.

The strength of this study was that we selected the concepts that seemed relevant to implementation research based on literature, modified them for electronic mental health settings, and improved the contents based on discussions with expert panels. Moreover, this study developed each questionnaire for users, providers, and managers or policy makers, all having an essential role in the implementation [8]. Evaluating the implementation outcomes of different stakeholders will clarify different perceptions of the intervention program, possibly leading to active knowledge exchange among users, providers, and consumers. Although our outcome measures need further evaluation, our study contributes to implementation research in digital mental health.

We acknowledge the following limitations of our study. First, it was vulnerable to selection bias. As we recruited participants via the internet for the psychometric validation study, they might not be representative of the general population in Japan. It is possible that the participants were more familiar with web-based programs, and they may have had a better understanding of digital mental health programs. In addition, this study conducted psychometric assessment for the outcome scales for users only because the intervention setting in which interested individuals enrolled themselves in the program did not involve any providers

or managers. In a study setting involving providers and managers or policy makers, the iOSDMH for providers and managers or policy makers will be needed to evaluate implementation outcomes. We plan to evaluate the iOSDMH for providers and managers or policy makers in our future intervention study (UMIN-CTR: ID UMIN 000036864). Another limitation is that criterion-related validity was not evaluated in the current psychometric assessment. The development process of the items may not be regarded as a theoretical approach. Future studies should evaluate criterion-related validity using other measures related to implementation concepts, such as the system usability scale or participation status of web-based programs. This study validated the Japanese version of the iOSDMH for users. Additional studies are needed for validating the English version. In future studies, we plan to apply these outcome measures in several web-based intervention trials to assess whether these implementation outcomes will predict the completion rate and participant attitude using digital access log information [35]. Although we tried to include multiple researchers in the digital mental health and implementation science domains from different countries, the iOSDMH scales would become more robust with a larger and more diverse review team. Finally, the setting in which we conducted the survey was an occupational setting (ie, for workers). Future studies should evaluate the scales in other settings (eg, clinical, school).

Conclusions

We developed implementation outcome scales for digital mental health interventions to assess the perceived outcomes for users, providers, and managers or policy makers. Psychometric assessment of the outcome scale for users showed acceptable reliability and validity. Future studies should apply the newly developed measures to assess the implementation status of the digital mental health program among different stakeholders and enhance collaborative problem-solving.

Acknowledgments

This work was supported by the Japan Society for the Promotion of Science under a Grant-in-Aid for Scientific Research (A) (grant 19H01073 to DN). This study received guidance from the National Center Consortium in Implementation Science for Health Equity (N-EQUITY) funded by the Japan Health Research Promotion Bureau (JH) Research Fund (grant 2019-(1)-4).

Authors' Contributions

DN was in charge of this project. NS and EO contributed to the development of the scale and conducted the survey. KI and NK ensured that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. NS and EO wrote the first draft of the manuscript, and all other authors revised the manuscript critically. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Implementation Outcome Scales for Digital Mental Health (iOSDMH).

[[DOCX File , 34 KB - formative v5i11e24332_app1.docx](#)]

References

1. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res* 2019 May;21(5):e13216 [FREE Full text] [doi: [10.2196/13216](https://doi.org/10.2196/13216)] [Medline: [31094356](https://pubmed.ncbi.nlm.nih.gov/31094356/)]
2. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR Ment Health* 2020 Mar;7(3):e18848 [FREE Full text] [doi: [10.2196/18848](https://doi.org/10.2196/18848)] [Medline: [32213476](https://pubmed.ncbi.nlm.nih.gov/32213476/)]
3. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. *Cogn Behav Ther* 2018 Jan;47(1):1-18. [doi: [10.1080/16506073.2017.1401115](https://doi.org/10.1080/16506073.2017.1401115)] [Medline: [29215315](https://pubmed.ncbi.nlm.nih.gov/29215315/)]
4. 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;74(4):351-359. [doi: [10.1001/jamapsychiatry.2017.0044](https://doi.org/10.1001/jamapsychiatry.2017.0044)] [Medline: [28241179](https://pubmed.ncbi.nlm.nih.gov/28241179/)]
5. Spek V, Cuijpers P, Nyklíček I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007 Mar;37(3):319-328. [doi: [10.1017/S0033291706008944](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]
6. Reins JA, Buntrock C, Zimmermann J, Grund S, Harrer M, Lehr D, et al. Efficacy and moderators of internet-based interventions in adults with subthreshold depression: an individual participant data meta-analysis of randomized controlled trials. *Psychother Psychosom* 2021;90(2):94-106 [FREE Full text] [doi: [10.1159/000507819](https://doi.org/10.1159/000507819)] [Medline: [32544912](https://pubmed.ncbi.nlm.nih.gov/32544912/)]
7. Aboujaoude E, Gega L, Parish MB, Hilty DM. Editorial: digital interventions in mental health: current status and future directions. *Front Psychiatry* 2020 Feb;11:111 [FREE Full text] [doi: [10.3389/fpsy.2020.00111](https://doi.org/10.3389/fpsy.2020.00111)] [Medline: [32174858](https://pubmed.ncbi.nlm.nih.gov/32174858/)]
8. A guide to implementation research in the prevention and control of noncommunicable diseases. World Health Organization. Geneva: World Health Organization; 2016. URL: <https://apps.who.int/iris/bitstream/handle/10665/252626/9789241511803-eng.pdf> [accessed 2020-12-01]
9. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. *J Clin Psychol* 2010 Apr;66(4):394-409 [FREE Full text] [doi: [10.1002/jclp.20659](https://doi.org/10.1002/jclp.20659)] [Medline: [20127795](https://pubmed.ncbi.nlm.nih.gov/20127795/)]
10. Hershenberg R, Satterthwaite TD, Daldal A, Katchmar N, Moore TM, Kable JW, et al. Diminished effort on a progressive ratio task in both unipolar and bipolar depression. *J Affect Disord* 2016 May;196:97-100 [FREE Full text] [doi: [10.1016/j.jad.2016.02.003](https://doi.org/10.1016/j.jad.2016.02.003)] [Medline: [26919058](https://pubmed.ncbi.nlm.nih.gov/26919058/)]
11. Coles ME, Ravid A, Gibb B, George-Denn D, Bronstein LR, McLeod S. Adolescent mental health literacy: young people's knowledge of depression and social anxiety disorder. *J Adolesc Health* 2016 Jan;58(1):57-62. [doi: [10.1016/j.jadohealth.2015.09.017](https://doi.org/10.1016/j.jadohealth.2015.09.017)] [Medline: [26707229](https://pubmed.ncbi.nlm.nih.gov/26707229/)]
12. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? a systematic review of quantitative and qualitative studies. *Psychol Med* 2015 Jan;45(1):11-27. [doi: [10.1017/S0033291714000129](https://doi.org/10.1017/S0033291714000129)] [Medline: [24569086](https://pubmed.ncbi.nlm.nih.gov/24569086/)]
13. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009 Apr;11(2):e13 [FREE Full text] [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
14. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *J Med Internet Res* 2004 Dec;6(4):e46 [FREE Full text] [doi: [10.2196/jmir.6.4.e46](https://doi.org/10.2196/jmir.6.4.e46)] [Medline: [15631970](https://pubmed.ncbi.nlm.nih.gov/15631970/)]
15. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a web-based self-help cognitive behavioral therapy program for panic disorder. *J Med Internet Res* 2005 Mar;7(1):e7 [FREE Full text] [doi: [10.2196/jmir.7.1.e7](https://doi.org/10.2196/jmir.7.1.e7)] [Medline: [15829479](https://pubmed.ncbi.nlm.nih.gov/15829479/)]
16. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011 Mar;38(2):65-76 [FREE Full text] [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
17. Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. *Adm Policy Ment Health* 2009 Jan;36(1):24-34 [FREE Full text] [Medline: [19104929](https://pubmed.ncbi.nlm.nih.gov/19104929/)]
18. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci* 2017 Aug;12(1):108 [FREE Full text] [doi: [10.1186/s13012-017-0635-3](https://doi.org/10.1186/s13012-017-0635-3)] [Medline: [28851459](https://pubmed.ncbi.nlm.nih.gov/28851459/)]
19. Lewis CC, Fischer S, Weiner BJ, Stanick C, Kim M, Martinez RG. Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. *Implement Sci* 2015 Nov;10:155 [FREE Full text] [doi: [10.1186/s13012-015-0342-x](https://doi.org/10.1186/s13012-015-0342-x)] [Medline: [26537706](https://pubmed.ncbi.nlm.nih.gov/26537706/)]
20. Hides L, Lubman DI, Elkins K, Catania LS, Rogers N. Feasibility and acceptability of a mental health screening tool and training programme in the youth alcohol and other drug (AOD) sector. *Drug Alcohol Rev* 2007 Sep;26(5):509-515. [doi: [10.1080/09595230701499126](https://doi.org/10.1080/09595230701499126)] [Medline: [17701514](https://pubmed.ncbi.nlm.nih.gov/17701514/)]

21. Moore GC, Benbasat I. Development of an instrument to measure the perceptions of adopting an information technology innovation. *Inf Syst Res* 1991 Sep;2(3):192-222. [doi: [10.1287/isre.2.3.192](https://doi.org/10.1287/isre.2.3.192)]
22. Whittingham K, Sofronoff K, Sheffield JK. Stepping Stones Triple P: a pilot study to evaluate acceptability of the program by parents of a child diagnosed with an autism spectrum disorder. *Res Dev Disabil* 2006 Aug;27(4):364-380. [doi: [10.1016/j.ridd.2005.05.003](https://doi.org/10.1016/j.ridd.2005.05.003)] [Medline: [16051461](https://pubmed.ncbi.nlm.nih.gov/16051461/)]
23. Yetter G. Assessing the acceptability of problem-solving procedures by school teams: preliminary development of the pre-referral intervention team inventory. *J Educ Psychol Consult* 2010 Jun;20(2):139-168. [doi: [10.1080/10474411003785370](https://doi.org/10.1080/10474411003785370)]
24. 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/)]
25. UTSMeD-Depression. URL: <http://www.utsumed-neo.xyz/> [accessed 2021-11-04]
26. Imamura K, Kawakami N, Tsuno K, Tsuchiya M, Shimada K, Namba K. Effects of web-based stress and depression literacy intervention on improving symptoms and knowledge of depression among workers: a randomized controlled trial. *J Affect Disord* 2016 Oct;203:30-37. [doi: [10.1016/j.jad.2016.05.045](https://doi.org/10.1016/j.jad.2016.05.045)] [Medline: [27280960](https://pubmed.ncbi.nlm.nih.gov/27280960/)]
27. Imamura K, Kawakami N, Tsuno K, Tsuchiya M, Shimada K, Namba K, et al. Effects of web-based stress and depression literacy intervention on improving work engagement among workers with low work engagement: an analysis of secondary outcome of a randomized controlled trial. *J Occup Health* 2017 Jan;59(1):46-54 [FREE Full text] [doi: [10.1539/joh.16-0187-OA](https://doi.org/10.1539/joh.16-0187-OA)] [Medline: [27885247](https://pubmed.ncbi.nlm.nih.gov/27885247/)]
28. Guttman L. Some necessary conditions for common-factor analysis. *Psychometrika* 1954 Jun;19(2):149-161. [doi: [10.1007/bf02289162](https://doi.org/10.1007/bf02289162)]
29. Maskey R, Fei J, Nguyen H. Use of exploratory factor analysis in maritime research. *Asian J Shipp Logist* 2018 Jun;34(2):91-111. [doi: [10.1016/j.ajsl.2018.06.006](https://doi.org/10.1016/j.ajsl.2018.06.006)]
30. Shumway M, Saunders T, Shern D, Pines E, Downs A, Burbine T, et al. Preferences for schizophrenia treatment outcomes among public policy makers, consumers, families, and providers. *Psychiatr Serv* 2003 Aug;54(8):1124-1128. [doi: [10.1176/appi.ps.54.8.1124](https://doi.org/10.1176/appi.ps.54.8.1124)] [Medline: [12883140](https://pubmed.ncbi.nlm.nih.gov/12883140/)]
31. 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;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/)]
32. Sidani S, Epstein DR, Bootzin RR, Moritz P, Miranda J. Assessment of preferences for treatment: validation of a measure. *Res Nurs Health* 2009 Aug;32(4):419-431 [FREE Full text] [doi: [10.1002/nur.20329](https://doi.org/10.1002/nur.20329)] [Medline: [19434647](https://pubmed.ncbi.nlm.nih.gov/19434647/)]
33. Hungerford DW, Pollock DA. Emergency department services for patients with alcohol problems: research directions. *Acad Emerg Med* 2003 Jan;10(1):79-84 [FREE Full text] [doi: [10.1111/j.1553-2712.2003.tb01982.x](https://doi.org/10.1111/j.1553-2712.2003.tb01982.x)] [Medline: [12511321](https://pubmed.ncbi.nlm.nih.gov/12511321/)]
34. Hungerford DW, Pollock DA, Todd KH. Acceptability of emergency department-based screening and brief intervention for alcohol problems. *Acad Emerg Med* 2000 Dec;7(12):1383-1392 [FREE Full text] [doi: [10.1111/j.1553-2712.2000.tb00496.x](https://doi.org/10.1111/j.1553-2712.2000.tb00496.x)] [Medline: [11099429](https://pubmed.ncbi.nlm.nih.gov/11099429/)]
35. Nishi D, Imamura K, Watanabe K, Obikane E, Sasaki N, Yasuma N, et al. Internet-based cognitive-behavioural therapy for prevention of depression during pregnancy and in the post partum (iPDP): a protocol for a large-scale randomised controlled trial. *BMJ Open* 2020 May;10(5):e036482 [FREE Full text] [doi: [10.1136/bmjopen-2019-036482](https://doi.org/10.1136/bmjopen-2019-036482)] [Medline: [32423941](https://pubmed.ncbi.nlm.nih.gov/32423941/)]

Abbreviations

EFA: exploratory factor analysis

iOSDMH: implementation outcome scales for digital mental health

Edited by G Eysenbach; submitted 15.09.20; peer-reviewed by A Wattanapisit, L Farrer; comments to author 24.11.20; revised version received 15.01.21; accepted 06.10.21; published 23.11.21.

Please cite as:

Sasaki N, Obikane E, Vedanthan R, Imamura K, Cuijpers P, Shimazu T, Kamada M, Kawakami N, Nishi D
Implementation Outcome Scales for Digital Mental Health (iOSDMH): Scale Development and Cross-sectional Study
JMIR Form Res 2021;5(11):e24332

URL: <https://formative.jmir.org/2021/11/e24332>

doi: [10.2196/24332](https://doi.org/10.2196/24332)

PMID: [34817391](https://pubmed.ncbi.nlm.nih.gov/34817391/)

©Natsu Sasaki, Erika Obikane, Rajesh Vedanthan, Kotaro Imamura, Pim Cuijpers, Taichi Shimazu, Masamitsu Kamada, Norito Kawakami, Daisuke Nishi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.11.2021. 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 Instruments for Reporting of Gastrointestinal Symptoms in Clinical Trials: Comparison of End-of-Day Diaries Versus the Experience Sampling Method

Abraham B Beckers¹, MD; Johanna T W Snijkers¹, MD; Zsa Zsa R M Weerts¹, MD, PhD; Lisa Vork¹, MD, PhD; Tim Klaassen¹, MD, PhD; Fabienne G M Smeets¹, MD, PhD; Ad A M Masclee¹, MD, Prof Dr; Daniel Keszthelyi¹, MD, PhD

Division of Gastroenterology-Hepatology, Department of Internal Medicine, Maastricht University Medical Center, Maastricht, Netherlands

Corresponding Author:

Abraham B Beckers, MD

Division of Gastroenterology-Hepatology

Department of Internal Medicine

Maastricht University Medical Center

Universiteitssingel 50

Maastricht, 6229 ER

Netherlands

Phone: 31 0433881844

Email: ab.beckers@maastrichtuniversity.nl

Abstract

Background: Questionnaires are necessary tools for assessing symptoms of disorders of the brain-gut interaction in clinical trials. We previously reported on the excellent adherence to a smartphone app used as symptom diary in a randomized clinical trial on irritable bowel syndrome (IBS). Other sampling methods, such as the experience sampling method (ESM), are better equipped to measure symptom variability over time and provide useful information regarding possible symptom triggers, and they are free of ecological and recall bias. The high frequency of measurements, however, could limit the feasibility of ESM in clinical trials.

Objective: This study aimed to compare the adherence rates of a smartphone-based end-of-day diary and ESM for symptom assessment in IBS and functional dyspepsia (FD).

Methods: Data from 4 separate studies were included. Patients with IBS participated in a randomized controlled trial, which involved a smartphone end-of-day diary for a 2+8-week (pretreatment + treatment) period, and an observational study in which patients completed ESM assessments using a smartphone app for 1 week. Patients with FD participated in a randomized controlled trial, which involved a smartphone end-of-day diary for a 2+12-week (pretreatment + treatment) period, and an observational study in which patients completed ESM assessments using a smartphone app for 1 week. Adherence rates were compared between these 2 symptom sampling methods.

Results: In total, 25 patients with IBS and 15 patients with FD were included. Overall adherence rates for the end-of-day diaries were significantly higher than those for ESM (IBS: 92.7% vs 69.8%, FD: 90.1% vs 61.4%, respectively).

Conclusions: This study demonstrates excellent adherence rates for smartphone app-based end-of-day diaries as used in 2 separate clinical trials. Overall adherence rates for ESM were significantly lower, rendering it more suitable for intermittent sampling periods rather than continuous sampling during longer clinical trials.

(*JMIR Form Res* 2021;5(11):e31678) doi:[10.2196/31678](https://doi.org/10.2196/31678)

KEYWORDS

irritable bowel syndrome; functional dyspepsia; digital diary; experience sampling method; smartphone app; mobile phone application; mHealth; eHealth; compliance; patient-reported outcome measures

Introduction

Disorders of brain-gut interaction (DBGI) are highly prevalent disorders, with a recent multinational survey study indicating that over 40% of the world's population has at least 1 DBGI [1]. Among the most common DBGI are irritable bowel syndrome (IBS) and functional dyspepsia (FD), which are characterized by lower and upper gastrointestinal symptoms, respectively, including abdominal pain, fullness, bloating, constipation, and diarrhea. Per definition, the diagnosis of these disorders is symptom-based, according to the Rome IV criteria [2]. By extension, the evaluation of treatment responses in clinical trials on DBGI relies completely on patient-reported outcome measures (PROMs). It is therefore of utmost importance that clinical trials use symptom sampling methods that are able to produce an accurate representation of the symptomatology as experienced by the patient. As such, paper symptom diaries have been scrutinized as they are prone to fake adherence, as subjects can fake or backfill written answers outside of the proposed time window to forge good adherence [3]. Thereby, the use of paper retrospective diaries introduces ecological and recall bias.

End-of-day symptom diaries are currently recommended by drug regulatory authorities to assess treatment response in IBS [4,5]. The widespread dissemination of the smartphone during the previous 2 decades creates possibilities for developing more advanced symptom sampling methods. We recently reported a digital end-of-day symptom diary using a smartphone app in a randomized controlled trial (RCT) on IBS [6]. We observed a very high adherence of almost 88% for the diary smartphone app during a treatment period of 8 weeks. End-of-day diaries in any form, however, are not free of the abovementioned ecological and recall biases owing to their retrospective nature. These limitations are currently best overcome by the experience sampling method (ESM), which employs random and repeated assessments at multiple time points across momentary states in daily life and thereby provides a detailed overview of symptoms experienced during the day. Previously developed ESM instruments for FD and IBS use a measurement frequency of 10 times a day [7,8]. It should be noted that the high frequency of this sampling method might raise concerns regarding adherence during clinical trials with a duration of several weeks or longer. In our previous IBS trial using the end-of-day diary,

adherence declined over time. Logging fatigue is considered the underlying cause of this decline in adherence. It could be hypothesized that this mechanism impairs adherence even more in methods with a higher sampling frequency, such as ESM. Hence, although ESM proved to be the method with more accurate real-life representation of daily symptoms, its feasibility in longer-duration clinical trials is still unknown. To draw conclusions regarding this question, a better understanding of ESM adherence and decline thereof is required.

In this exploratory study, we sought to compare adherence for end-of-day diaries used in 2 RCTs with adherence for ESM in 2 separate observational studies. We hypothesized that overall adherence would be superior with the use of the end-of-day diary, as compared to ESM. Moreover, we hypothesized that adherence would remain more stable over time for the end-of-day diary than for ESM.

Methods

Methods Overview

This study is based on data from 2 RCTs and 2 observational studies. For each study type, 1 study focused on IBS and 1 on FD. The RCTs used end-of-day diaries, whereas the observational studies used momentary assessments (ESM). The Rome IV criteria were used as inclusion criteria in each study in accordance with the disorder being investigated. All 4 studies had been approved by the Maastricht University Medical Center+ (MUMC+) ethics committee. All study procedures were performed in compliance with Good Clinical Practice Guidelines and in accordance with the revised Declaration of Helsinki. All subjects gave written informed consent prior to participation.

Although including data from multiple studies, all patients included in this study were required to have participated in both an RCT and an observational study (on the same disorder) to reduce variability across individuals. All patients thereby completed both end-of-day diaries during a longer period of time and ESM for a period of 1 week (not simultaneously). Details on each study are provided below, and an overview of sampling characteristics is presented in [Table 1](#). The exact queries in each sampling method are provided in [Multimedia Appendix 1](#).

Table 1. Overview of sampling specifics per study.

Sampling specifics	Randomized controlled trial on irritable bowel syndrome	Observational study on irritable bowel syndrome	Randomized controlled trial on functional dyspepsia	Observational study on functional dyspepsia
Sampling method	Digital end-of-day diary (smartphone app)	End-of-day paper diary and digital experience sampling method (smartphone app) ^a	Digital end-of-day diary (smartphone app)	Combined digital end-of-day diary and experience sampling method (smartphone app)
Sampling duration	2 weeks pretreatment + 8 weeks treatment	1 week	2 weeks pretreatment plus 12 weeks treatment	1 week
Sampling frequency	Once a day	10 times per day	Once a day	10 times per day
Sampling timeframe	Between 6 PM and 12 PM	Between 7 AM and 10 PM	Between 7 PM and 12 PM	Between 7 AM and 10 PM
Push notifications	Once at 10 PM	Randomly timed	Once at 9 PM	Randomly timed
Items, n	3 ^b	32 (5 domains)	7 ^c	33 (4 domains)
Estimated time investment	15-30 seconds	2-3 minutes (per assessment)	15-30 seconds	2-3 minutes (per assessment)

^aNote that the end-of-day diaries are similar to the ones used in the corresponding randomized controlled trial. In the observational study, however, end-of-day diaries were completed on paper.

^bIn addition to items in the Bristol stool chart and adverse event and sporadic medication use queries.

^cIn addition to adverse event and sporadic medication use queries.

RCTs Using End-of-Day Diaries

The RCT on IBS has been discussed in detail elsewhere [9]. In brief, the primary aim was to investigate the efficacy of peppermint oil, a conventional small-intestinal and a novel ileocolonic release formulation, in patients with IBS (ClinicalTrials.gov, NCT02716285). In this randomized, placebo-controlled trial, patients with IBS aged 18-75 years were included. Patients completed an end-of-day diary using a smartphone app during a pretreatment period (2 weeks) and treatment period (8 weeks), as described previously [6]. At the core, this diary consisted of 1 question regarding abdominal pain experienced each day (to be scored on an 11-point numerical scale). After completing the abdominal pain question, subjects were asked about adverse events and sporadic medication use. During the day, patients had the option to report on defecation in accordance with the Bristol stool chart [10]. Patients were instructed to register abdominal pain daily between 6 PM and 12 PM. Finally, psychological comorbidities were assessed at baseline using the General Anxiety Disorder, 7-item (GAD-7) scale and Patient Health Questionnaire, 9-item (PHQ-9).

The second RCT is an ongoing trial that investigates the efficacy of nortriptyline in patients with FD (ClinicalTrials.gov, NCT03652571). In this randomized, placebo-controlled trial, patients with FD aged 18-65 years were included. Patients completed an end-of-day diary using a smartphone app during a pretreatment period (2 weeks) and treatment period (12 weeks); the app was similar to the one used in the RCT on IBS. In the RCT on FD, however, the diary consisted of questions corresponding to the five core symptoms of FD: epigastric pain, epigastric burning, early satiety, postprandial fullness, and upper abdominal bloating [11]. In addition to these 5 questions, subjects were asked about adverse events and sporadic medication use. Patients are instructed to register symptoms daily between 7 PM and 12 PM. There was no registration of bowel movements in this trial, as an altered bowel habit is not

a core symptom in FD. Finally, psychological comorbidities were assessed at baseline using the GAD-7 scale and PHQ-9.

Observational Studies Using ESM

ESM data from patients with IBS were obtained from a validation study of a newly developed patient-reported outcome measure (based on the ESM) for the use in IBS (ClinicalTrials.gov, NCT02880722) [8]. Patients with IBS and healthy volunteers between 18 and 70 years of age were included in this study. Both groups completed an end-of-day paper diary and ESM for a period of 1 week. The ESM was incorporated in a customized smartphone app. The ESM consisted of 10 assessments randomly timed between 7:30 AM and 10:30 PM. Each assessment was precluded by an auditory signal, and the app was programmed to enable completion of the assessment within 10 minutes after the auditory signal. Subjects were instructed to complete as many assessments as possible, but to pass over questionnaires when completing was not feasible (eg, when driving). Assessments covered five different domains, as described previously [8]: physical status (eg, abdominal pain), defecation (since the previous auditory signal), psychological factors (eg, positive and negative affect), environment (eg, current location and company), and nutrition and drug use. In total, the ESM for use in IBS consisted of 32 items (per assessment).

ESM data from patients with FD were obtained from a separate validation study of a newly developed patient-reported outcome measure (based on the ESM) for the use in FD (ClinicalTrials.gov, NCT04204421) [12]. Patients with FD and healthy volunteers between 18 and 75 years of age were included in this study. Both groups completed an end-of-day diary and ESM for a period of 1 week. The diary and ESM were incorporated in the same customized smartphone app. ESM was used in a manner similar to the IBS ESM study, with 10 assessments randomly timed between 7 AM and 10 PM. Assessments in the FD ESM study covered 4 domains, which

included the same as those in the IBS ESM study [7]. In total, the ESM for use in FD consisted of 33 items (per assessment).

Statistical Analyses and Data Plots

Adherence was the primary outcome measure. For both the end-of-day diaries and ESM, overall adherence was calculated as the percentage of completed assessments throughout the study. For visualizing adherence over time, weekly adherence rates were plotted for the clinical trials (end-of-day diaries) and daily adherence rates for the observational studies (ESM). In addition, for ESM we also calculated overall adherence as the number of days on which ≥ 6 of the 10 assessments were completed, as described previously [8,13]. The latter can be considered more appropriate when evaluating adherence of sampling methods such as ESM, where an excess of measurements is provided to obtain sufficient data during the day [13,14].

All data were plotted using MATLAB R2018a. Linear mixed models were performed using the lme4 function in R Statistical Software (version 3.6.3, February 29, 2020) [15]. In each model (per study), adherence to the app constituted the dependent variable and time constituted the within-subject independent variable. A restricted maximum likelihood estimation method and first-order autoregressive variance-covariance matrix for the within-subject variable time fitted the data best on the basis of the lowest value of the Akaike information criterion.

Given the exploratory nature of the study, we did not perform sample size calculations.

Results

In total, 25 patients with IBS and 15 patients with FD were included in our analysis for adherence comparison. An overview

of subject characteristics is provided in Table 2. Both patients with IBS and those with FD were more frequently female (72.0% and 73.3%, respectively). For patients with FD, the time between participation in the 2 studies was significantly longer than that for patients with IBS (12.5 vs 7.6 months, respectively; $t_{26,12}=4.30$; $P<.001$). All subjects in the IBS and FD studies participated in the ESM observational study after participating in the RCT.

Adherence rates for the end-of-day diary in the IBS RCT during the pretreatment period, treatment period, and total study duration (both periods combined) were 93.4%, 92.6%, and 92.7%, respectively. Overall adherence—ie, the percentage of total completed assessments—for ESM during the observational IBS study was 69.8%.

Adherence rates for the end-of-day diary in the FD RCT during the pretreatment period, treatment period, and total study duration were 92.9%, 89.7%, and 90.1%, respectively. Overall adherence for ESM during the observational FD study was 61.4%.

Of note, for trials using the ESM method, completion of ≥ 6 of the 10 questionnaires per day is considered as being adherent, as described previously [8,13]. This type of adherence calculation can be considered more representative for sampling methods such as ESM. When using this approach, overall adherence was 79.4% and 64.8% for the ESM IBS and ESM FD studies, respectively. Adherence in the latter was noticeably lower owing to the effect of 4 outliers (adherence < 15%) in this relatively small group. Three of 4 subjects reported a specific reason for low adherence, which included (1) a technical error (subject did not receive push notifications on his/her smartphone), (2) attending the funeral of a close relative, and (3) not being able to complete most assessments during day job.

Table 2. Summary of patient demographic and baseline characteristics.

Characteristics	Randomized controlled trial on irritable bowel syndrome	Observational study on irritable bowel syndrome	Randomized controlled trial on functional dyspepsia	Observational study on functional dyspepsia
Subjects, n	25		15	
Time between study participations (months), mean (SD); range	7.6 (3.1); 1-17		12.5 (3.6); 7-18	
Age ^a (years), mean (SD); range	35.9 (12.8); 22-59		41.4 (15.2); 18-64	
Females, n (%)	18 (72.0)		11 (73.3)	
Educational level, n (%)				
No education	0		0	
Low	2 (8.0%)		1 (6.7%)	
Moderate	12 (48.0%)		5 (33.3%)	
High	11 (44.0%)		9 (60.0%)	
Irritable bowel syndrome or functional dyspepsia subtype, n (%)	<ul style="list-style-type: none"> • Diarrhea: 14 (56.0) • Constipation: 3 (12.0) • Mixed: 4 (16.0) • Undefined: 4 (16.0) 		<ul style="list-style-type: none"> • Postprandial distress: 5 (33.3) • Epigastric pain: 4 (26.7) • Overlap: 6 (40.0) 	
Irritable bowel syndrome or functional dyspepsia severity^b				
Mean score (SD)	228.8 (24.5)		83.2 (22.8)	
Mild, n (%)	7 (28.0)		— ^c	
Moderate, n (%)	13 (52.0)		—	
Severe, n (%)	5 (20.0)		—	
Psychological comorbidities^d				
Anxiety				
Score, mean (SD)	4.2 (2.9)		3.3 (2.8)	
Minimal, n (%)	14 (56.0)		12 (80.0)	
Mild, n (%)	9 (36.0)		2 (13.3)	
Moderate, n (%)	2 (8.0)		1 (6.7)	
Depression				
Score, mean (SD)	5.0 (2.7)		4.9 (4.2)	
Minimal, n (%)	14 (56.0)		8 (53.3)	
Mild, n (%)	10 (40.0)		6 (40)	
Moderate, n (%)	1 (4.0)		1 (6.7)	

^aAge upon registering for the first study.

^bFor irritable bowel syndrome symptom severity, the Irritable Bowel Syndrome Severity Scoring System (IBS-SSS) was used. Scores were defined as follows: <175, mild; 175-300, moderate; and >300, severe. For functional dyspepsia symptom severity, the Nepean Dyspepsia Index (NDI symptom scale) was used (continuous scale only, no validated severity categories).

^c—: not determined.

^dFor anxiety, the General Anxiety Disorder scale, 7-item, was used, and for depression, the Patient Health Questionnaire, 9-item, was used. Scores were defined as follows: ≥5, mild; ≥10, moderate; and ≥15, severe.

Weekly adherence to the end-of-day diaries and daily adherence to the ESM app are shown in [Figure 1](#). Linear mixed models did not demonstrate a significant decline in adherence over time in either RCT (main effect of time; IBS RCT: $F_{1,224}=2.24$; $P=.14$; FD RCT: $F_{1,194}=0.87$; $P=.77$). A minor decline in adherence over time can be observed in the lower panels corresponding to the ESM studies, which was not significant

(linear mixed models, main effect of time; ESM IBS study: $F_{1,149}=3.41$; $P=.07$; ESM FD study: $F_{1,89}=1.23$; $P=.27$).

Cumulative completed assessments are plotted against the total number of assessments for each study and each subject in [Figures 2-5](#). Single subject plots can be compared as subplot positions in the figures correspond to the same subject.

Figure 1. Adherence to each symptom assessment app. For end-of-day diaries, the weekly adherence is shown (ie, percentage of completed assessments for each week). For the experience sampling method, the daily adherence is shown (ie, percentage of completed assessments for each day [out of 10 measurements]). ESM: experience sampling method, FD: functional dyspepsia, IBS: irritable bowel syndrome.

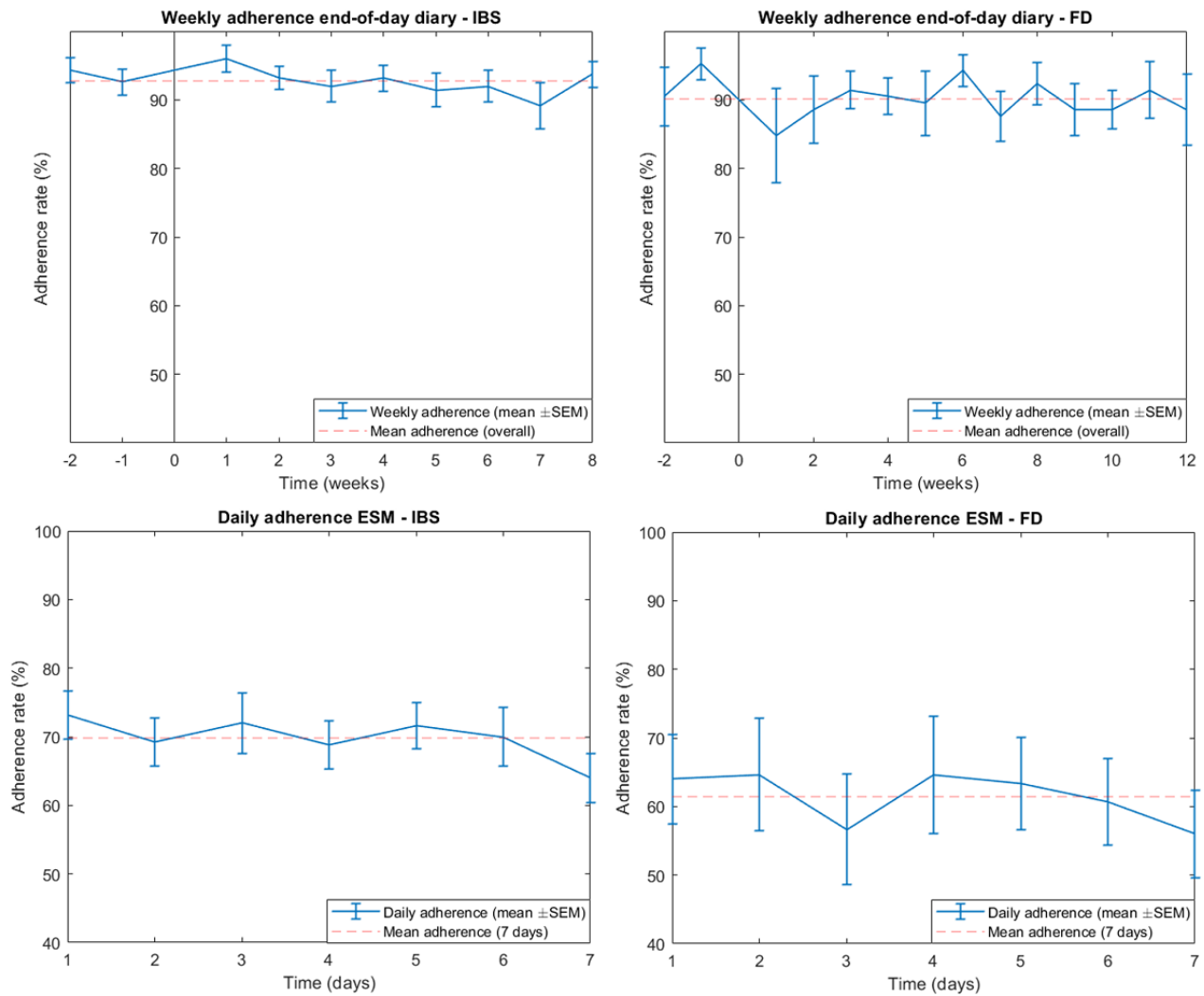


Figure 2. Cumulative completed diary assessments in the randomized controlled trial on irritable bowel syndrome. The red line indicates maximum number of assessments, and the blue line indicates actual number of completed assessments. Note that subplot positions in this figure correspond to the same subject in [Figure 3](#). X-axis: assessment number, Y-axis: completed assessments.

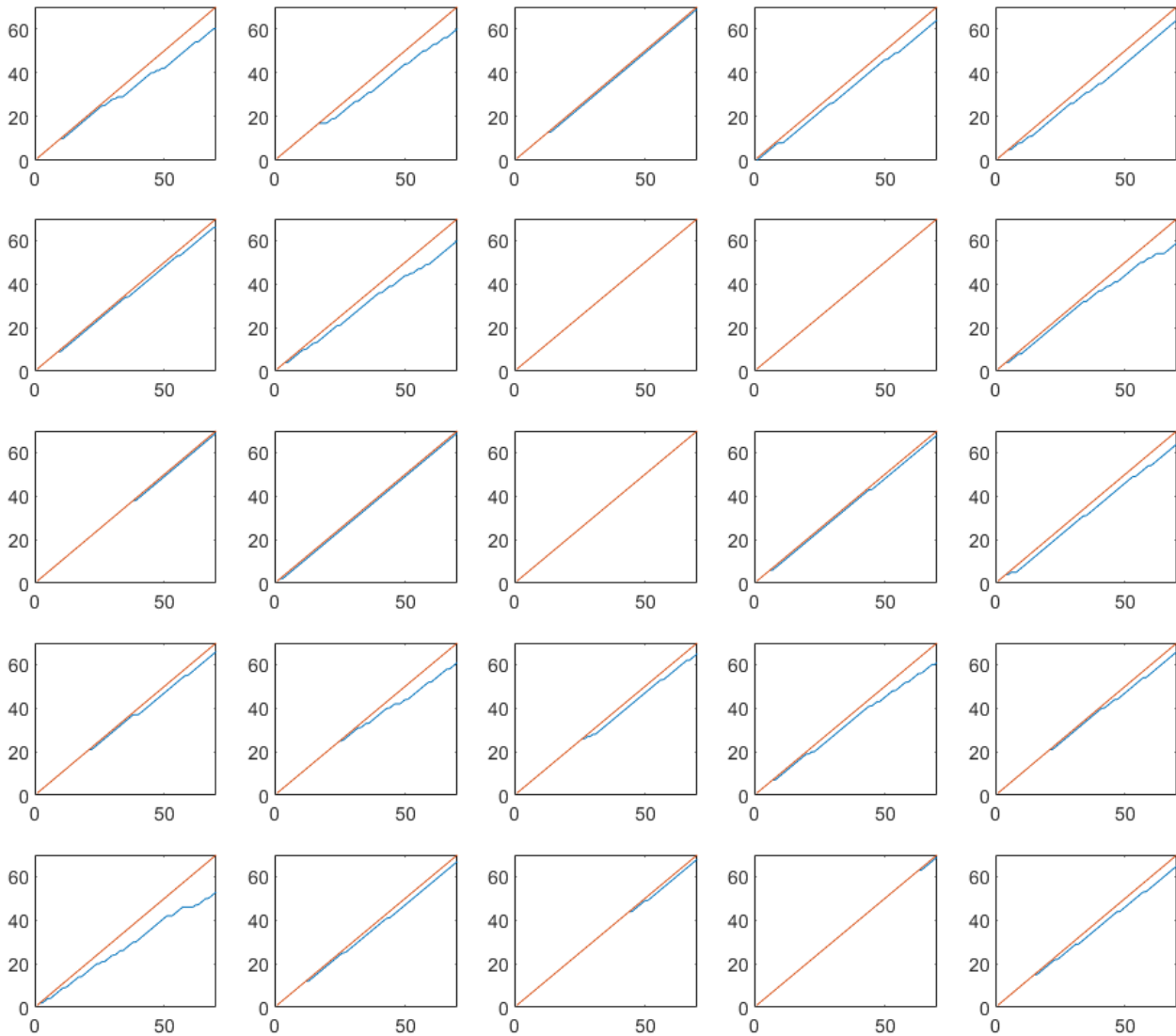


Figure 3. Cumulative completed experience sampling method assessments in the observational study on irritable bowel syndrome. The red line indicates maximum number of assessments, and the blue line indicates actual number of completed assessments. Note that subplot positions in this figure correspond to the same subject in [Figure 2](#). X-axis: assessment number, Y-axis: completed assessments.

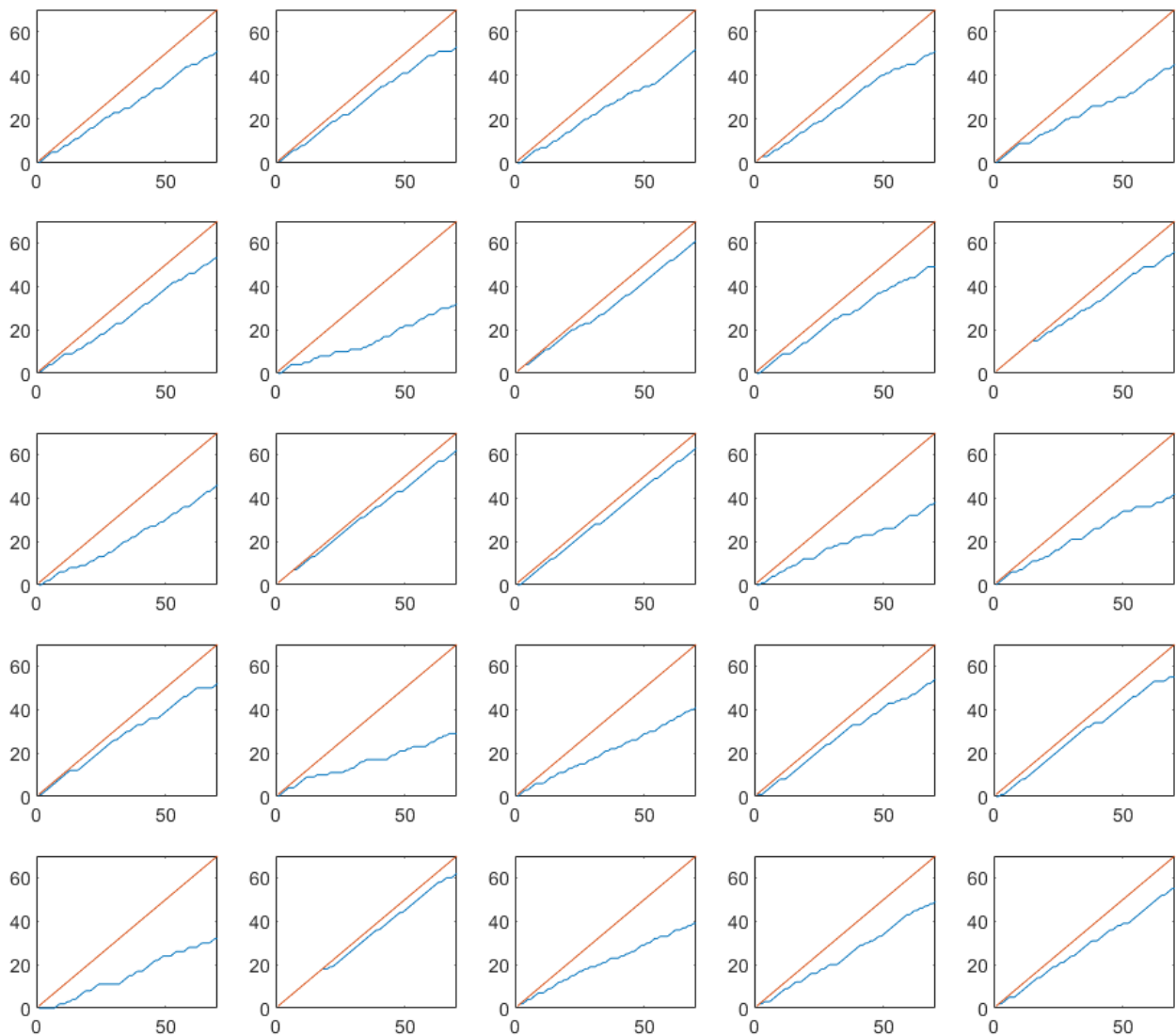


Figure 4. Cumulative completed diary assessments in the randomized controlled trial on functional dyspepsia. The red line indicates maximum number of assessments, and the blue line indicates actual number of completed assessments. Note that subplot positions in this figure correspond to the same subject in Figure 5. X-axis: assessment number, Y-axis: completed assessments.

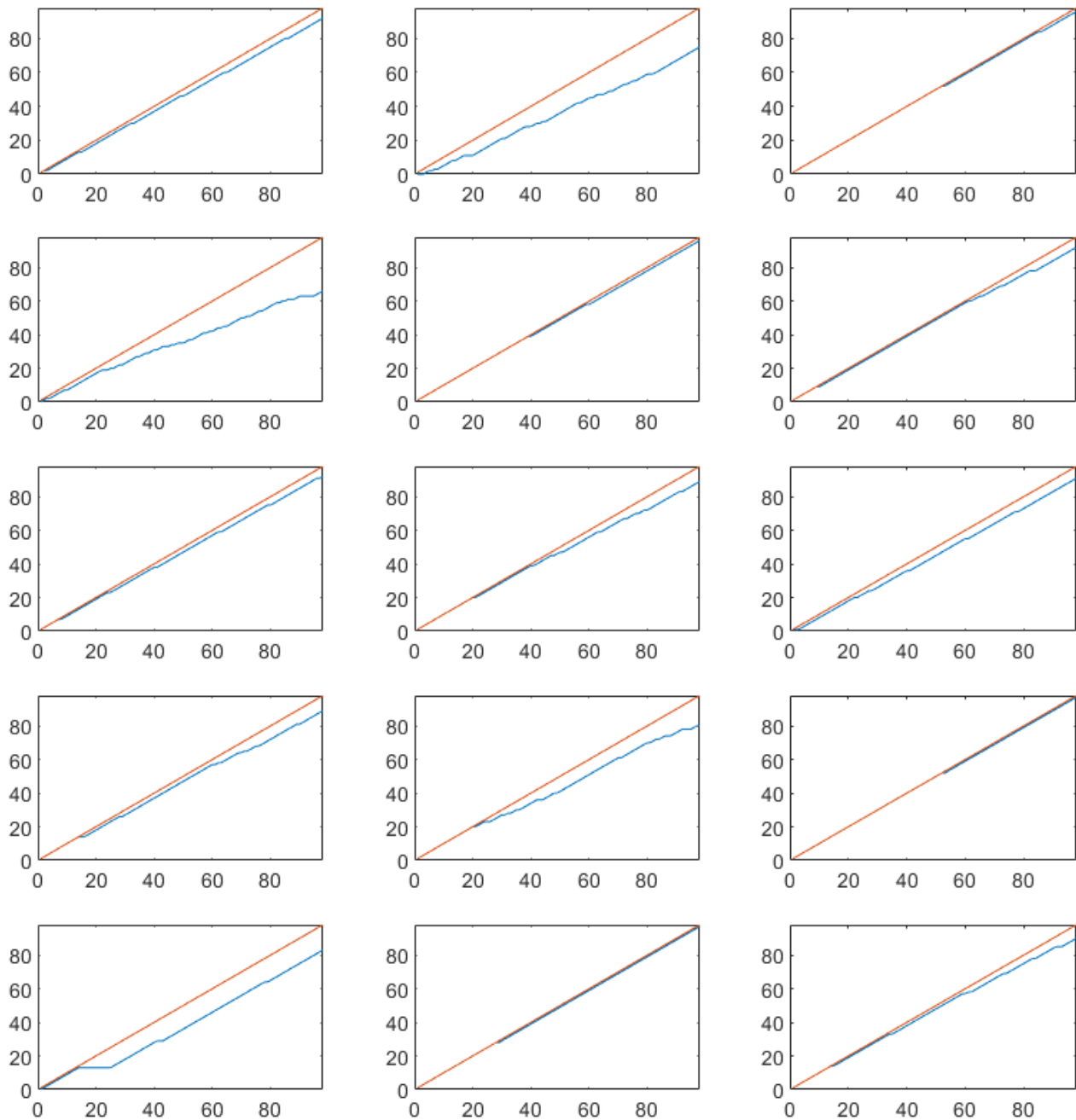
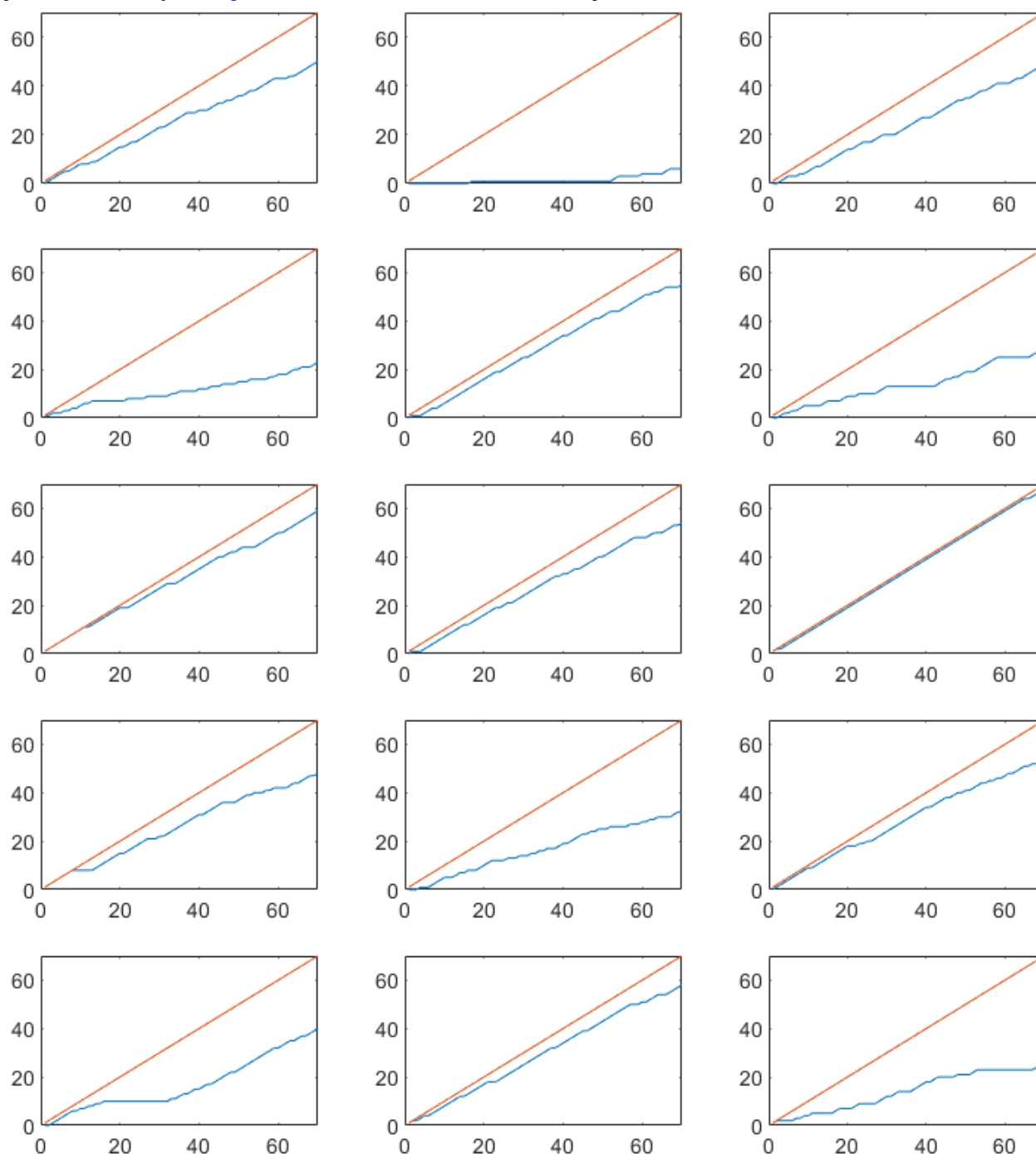


Figure 5. Cumulative completed experience sampling method assessments in the observational study on functional dyspepsia. The red line indicates maximum number of assessments, and the blue line indicates actual number of completed assessments. Note that subplot positions in this figure correspond to the same subject in Figure 4. X-axis: assessment number, Y-axis: completed assessments.



Discussion

Principal Findings

In this study, we explored adherence rates for smartphone apps used for symptom assessment in IBS and FD. In line with our previous findings in a large RCT on IBS [6], we found an excellent overall adherence of 90% for a digital end-of-day diary in an ongoing trial with patients with FD. Given that these diaries only enable the logging of symptoms experienced during the day that the diary is being filled in, fake adherence (ie, backfilling) is completely prevented. Indeed, the FDA recommends daily symptom assessment in IBS trials, which is

best facilitated by the digital (smartphone) framework presented here [4]. The high overall adherence rates as observed in the FD and in the IBS RCTs confirm the feasibility of these digital diaries in clinical trials. Importantly, the RCT in FD involves 5 diary questions as opposed to a single question in the diary used in the IBS RCT. In addition, the FD RCT is 4 weeks longer in duration than the IBS RCT. It is encouraging that regardless of the added burden, adherence for the end-of-day diary in the FD RCT is still excellent.

In our previous IBS RCT where we included 189 patients, we reported a small but significant decrease in adherence for the completion of daily diaries over the study duration. Such a

decrease in adherence can be referred to as logging fatigue. In the current study, we found no evidence of logging fatigue in the subset of the IBS RCT or in the FD RCT. We hypothesized that with more frequent assessments, such as with ESM, logging fatigue could become an increasing issue. Interestingly, in the current study, we found no significant decline in adherence rates over time in both ESM studies. However, since both studies were of relatively short duration (7 days), we cannot draw any conclusions on possible declining adherence rates over time when ESM is used for longer periods.

Overall adherence rates for the ESM studies were evidently lower than the end-of-day diary adherence rates, though in line with rates reported in previous studies [13,16]. Even when considering that it is generally not feasible to complete all ESM assessments and calculating ESM adherence as the number of days that ≥ 6 of the 10 assessments are completed, ESM adherence rates were still noticeably lower. To a large extent, this will likely be related to the nature of measurements. For the end-of-day diaries, subjects can choose a suitable moment between 6 PM (or 7 PM) and 12PM, and can do so each day as per their own schedule. For ESM, on the other hand, measurements are by their very definition timed at random moments and should be completed within 10 minutes after the assessment was announced. ESM is, therefore, likely to involve measurements at times when the subject is not able to complete the symptom assessment, especially as there will always be measurements within working hours. Furthermore, it is easy to miss a haptic or auditory signal on your cellphone. Moreover, the ESM assessments were far more extensive than the end-of-day diaries as they involved multiple domains; for example, physical, psychological, environmental, and nutritional domains. Indeed, it was demonstrated in a systematic review of studies including electronic diaries of various lengths that the extent of the diary used was negatively associated with adherence [16]. The large difference in overall adherence rates between ESM and end-of-day diaries may reflect on the higher burden of ESM. Therefore, we think that it is not feasible to use ESM continuously during a trial of several weeks, especially because adherence already tends to be lower in studies of longer duration [17]. A solution could be to use ESM intermittently (eg, 1 week in every 4 weeks), complementing the end-of-day diary. Thus, the end-of-day diary provides a strong continuous measurement framework, where ESM can be used at fixed periods to examine changes in symptom variability and symptom triggers over time (ie, during treatment), in addition to analyzing the complexity of factors contributing to symptom perception. However, the responsiveness (ie, the sensitivity to detect change over time) of ESM has not yet been evaluated.

Finally, one should appreciate the differences in acquired data when using ESM or the end-of-day diary. As already mentioned above, ESM provides more detailed information on symptoms and their possible triggers. Our preference for the end-of-day diary as a continuous measurement framework primarily relates to clinical trials, as this is also in accordance with current recommendations from regulatory authorities. It is possible that in some situations the more detailed data outweighs the drawback of the higher number of missing values. This could especially be the case in clinical practice of functional disorders, where additional information on symptom triggers is extremely valuable.

Limitations

A limitation of the current study is the small size of the study population. This is mitigated by the within-subject nature of the study, as the obligatory participation in both an end-of-day diary and ESM study limits subject specific effects on adherence, aiding a better comparison of assessment methods. As mentioned above, the possibility of selection bias cannot be excluded, as subjects who are willing to participate in more than one study could have a very strong motivation, which may translate to unrepresentatively high adherence rates. On the other hand, the overall adherence rate of our subset of subjects from the IBS RCT was only a few percentage points higher than that of the whole group, arguing against such selection bias. Finally, since all subjects participated in the ESM studies after participation in an RCT, a carry-over treatment effect could have affected logging adherence. However, this would likely have influenced adherence during the RCT itself as well, and we observed stable adherence during both RCTs. Moreover, we previously observed no effect of GI symptoms on adherence rates. The latter also suggests that variation in duration between ESM and RCT participation is less relevant.

Conclusions

In conclusion, we here demonstrate excellent adherence rates for smartphone app-based end-of-day diaries as well as good adherence to 2 ESM-based apps. Overall adherence rates for ESM were evidently lower, as would be expected given the nature of the methodology, but possibly also reflecting on the larger burden of this sampling method given the higher number of cues and questions to be answered. Even though we could not demonstrate a decline in response rate with ESM over a period of 7 days, it seems unfeasible to use ESM continuously in clinical trials over several weeks. Given the added value of ESM, however, researchers should consider complementing end-of-day diaries with intermittent periods of ESM.

Acknowledgments

No specific funding was received for the current study. The irritable bowel syndrome (IBS) randomized controlled trial (RTC) was funded by ZonMw (The Netherlands Organization for Health Research and Development) with co-funding from Will Pharma. The IBS observational study was funded in part by Grunenthal. The functional dyspepsia (FD) RCT is funded by ZonMw. No specific funding was received for the FD observational trial.

Authors' Contributions

ABB conceptualized and designed the study and drafted the manuscript. ABB, JTWS, ZZRMW, LV, TK, and FGMS collected the data. ABB, JTWS, ZZRMW, LV, TK, and FGMS cleaned the data. ABB, JTWS, AAMM, and DK analyzed and interpreted the data. JTWS, ZZRMW, LV, TK, FGMS, AAMM, and DK constructively reviewed the manuscript. AAMM and DK obtained funding. All authors approved the final manuscript.

Conflicts of Interest

AAMM and DK have received an unrestricted research grant from Will Pharma, which also supported ZZRMW and LV in attending a scientific meeting. AAMM and DK have received funding from Allergan, ZonMw, Grunenthal. DK has received funding from the Dutch Foundation for Gastroenterology-Hepatology (MLDS), United Europe Gastroenterology (UEG), Rome Foundation, and EU Horizon 2020, outside of submitted work. AAMM has received funding from the Dutch Foundation for Cancer Research (KWF), Pentax, outside of submitted work. TK received a salary from Will Pharma as part of the "Subsidie MKB Innovatiestimulerend Topsectoren" (MIT), outside of submitted work. ABB, JTWS, and FGMS have nothing to declare.

Multimedia Appendix 1

Supplementary Information.

[[DOCX File, 20 KB - formative_v5i11e31678_app1.docx](#)]

References

1. Sperber AD, Bangdiwala SI, Drossman DA, Ghoshal UC, Simren M, Tack J, et al. Worldwide prevalence and burden of functional gastrointestinal disorders, results of Rome Foundation global study. *Gastroenterology* 2021 Jan;160(1):99-114.e3 [FREE Full text] [doi: [10.1053/j.gastro.2020.04.014](https://doi.org/10.1053/j.gastro.2020.04.014)] [Medline: [32294476](https://pubmed.ncbi.nlm.nih.gov/32294476/)]
2. Drossman DA. Functional gastrointestinal disorders: history, pathophysiology, clinical features and Rome IV. *Gastroenterology* 2016 Mar 19:00223-00227. [doi: [10.1053/j.gastro.2016.02.032](https://doi.org/10.1053/j.gastro.2016.02.032)] [Medline: [27144617](https://pubmed.ncbi.nlm.nih.gov/27144617/)]
3. Lauritsen K, Degl' Innocenti A, Hendel L, Praest J, Lytje MF, Clemmensen-Rotne K, et al. Symptom recording in a randomised clinical trial: paper diaries vs. electronic or telephone data capture. *Control Clin Trials* 2004 Dec;25(6):585-597. [doi: [10.1016/j.cct.2004.09.001](https://doi.org/10.1016/j.cct.2004.09.001)] [Medline: [15588745](https://pubmed.ncbi.nlm.nih.gov/15588745/)]
4. D'Agostino RB, Sullivan L, Massaro J, editors. Center for Drug Evaluation and Research (CDER). In: *Wiley Encyclopedia of Clinical Trials*. Hoboken, NJ: John Wiley & Sons, Inc; 2008.
5. Guideline on the evaluation of medicinal products for the treatment of irritable bowel syndrome. European Medicines Agency. 2014. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-medicinal-products-treatment-irritable-bowel-syndrome-revision-1_en.pdf [accessed 2021-09-29]
6. Weerts ZZRM, Heinen KGE, Masclee AAM, Qanjel ABA, Winkens B, Vork L, et al. Smart data collection for the assessment of treatment effects in irritable bowel syndrome: observational study. *JMIR Mhealth Uhealth* 2020 Nov 02;8(11):e19696 [FREE Full text] [doi: [10.2196/19696](https://doi.org/10.2196/19696)] [Medline: [33030150](https://pubmed.ncbi.nlm.nih.gov/33030150/)]
7. Smeets FGM, Keszthelyi D, Vork L, Tack J, Talley NJ, Simren M, et al. Development of a real-time patient-reported outcome measure for symptom assessment in patients with functional dyspepsia using the experience sampling method. *Neurogastroenterol Motil* 2019 Feb;31(2):e13496. [doi: [10.1111/nmo.13496](https://doi.org/10.1111/nmo.13496)] [Medline: [30393939](https://pubmed.ncbi.nlm.nih.gov/30393939/)]
8. Vork L, Keszthelyi D, Mujagic Z, Kruijmel JW, Leue C, Pontén I, et al. Development, content validity, and cross-cultural adaptation of a patient-reported outcome measure for real-time symptom assessment in irritable bowel syndrome. *Neurogastroenterol Motil* 2018 Mar;30(3). [doi: [10.1111/nmo.13244](https://doi.org/10.1111/nmo.13244)] [Medline: [29106029](https://pubmed.ncbi.nlm.nih.gov/29106029/)]
9. Weerts ZZRM, Masclee AAM, Witteman BJM, Clemens CHM, Winkens B, Brouwers JRBJ, et al. Efficacy and safety of peppermint oil in a randomized, double-blind trial of patients with irritable bowel syndrome. *Gastroenterology* 2020 Jan;158(1):123-136. [doi: [10.1053/j.gastro.2019.08.026](https://doi.org/10.1053/j.gastro.2019.08.026)] [Medline: [31470006](https://pubmed.ncbi.nlm.nih.gov/31470006/)]
10. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol* 1997 Sep;32(9):920-924. [doi: [10.3109/00365529709011203](https://doi.org/10.3109/00365529709011203)] [Medline: [9299672](https://pubmed.ncbi.nlm.nih.gov/9299672/)]
11. Taylor F, Higgins S, Carson RT, Eremenco S, Foley C, Lacy BE, Patient-Reported Outcome Consortium s Functional Dyspepsia Working Group. Development of a symptom-focused patient-reported outcome measure for functional dyspepsia: The Functional Dyspepsia Symptom Diary (FSDS). *Am J Gastroenterol* 2018 Jan;113(1):39-48 [FREE Full text] [doi: [10.1038/ajg.2017.265](https://doi.org/10.1038/ajg.2017.265)] [Medline: [28925989](https://pubmed.ncbi.nlm.nih.gov/28925989/)]
12. Klaassen T, Smeets FGM, Vork L, Tack J, Talley NJ, Simren M, et al. Psychometric evaluation of an experience sampling method-based patient-reported outcome measure in functional dyspepsia. *Neurogastroenterol Motil* 2021 May 02:e14136. [doi: [10.1111/nmo.14136](https://doi.org/10.1111/nmo.14136)] [Medline: [33934444](https://pubmed.ncbi.nlm.nih.gov/33934444/)]
13. Mujagic Z, Leue C, Vork L, Lousberg R, Jonkers DMAE, Keszthelyi D, et al. The Experience Sampling Method--a new digital tool for momentary symptom assessment in IBS: an exploratory study. *Neurogastroenterol Motil* 2015 Sep;27(9):1295-1302. [doi: [10.1111/nmo.12624](https://doi.org/10.1111/nmo.12624)] [Medline: [26100684](https://pubmed.ncbi.nlm.nih.gov/26100684/)]

14. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. *Psychol Assess* 2012 Sep;24(3):713-720. [doi: [10.1037/a0026733](https://doi.org/10.1037/a0026733)] [Medline: [22250597](https://pubmed.ncbi.nlm.nih.gov/22250597/)]
15. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015;67(1):1-48. [doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)]
16. Morren M, van Dulmen S, Ouwerkerk J, Bensing J. Compliance with momentary pain measurement using electronic diaries: a systematic review. *Eur J Pain* 2009 Apr;13(4):354-365. [doi: [10.1016/j.ejpain.2008.05.010](https://doi.org/10.1016/j.ejpain.2008.05.010)] [Medline: [18603458](https://pubmed.ncbi.nlm.nih.gov/18603458/)]
17. Burton C, Weller D, Sharpe M. Are electronic diaries useful for symptoms research? A systematic review. *J Psychosom Res* 2007 May;62(5):553-561. [doi: [10.1016/j.jpsychores.2006.12.022](https://doi.org/10.1016/j.jpsychores.2006.12.022)] [Medline: [17467410](https://pubmed.ncbi.nlm.nih.gov/17467410/)]

Abbreviations

DBGI: disorders of brain-gut interaction
ESM: experience sampling method
FD: functional dyspepsia
GAD-7: General Anxiety Disorder, 7-item
IBS: irritable bowel syndrome
IBS-SSS: Irritable Bowel Syndrome Severity Scoring System
MUMC+: Maastricht University Medical Center+
NDI: Nepean Dyspepsia Index
PHQ-9: Patient Health Questionnaire, 9-item
PROM: patient-reported treatment measure
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 30.06.21; peer-reviewed by L Guadagnoli, J Edwards; comments to author 28.07.21; revised version received 09.08.21; accepted 17.09.21; published 24.11.21.

Please cite as:

Beckers AB, Snijkers JTW, Weerts ZZRM, Vork L, Klaassen T, Smeets FGM, Masclee AAM, Keszthelyi D
Digital Instruments for Reporting of Gastrointestinal Symptoms in Clinical Trials: Comparison of End-of-Day Diaries Versus the Experience Sampling Method
JMIR Form Res 2021;5(11):e31678
URL: <https://formative.jmir.org/2021/11/e31678>
doi: [10.2196/31678](https://doi.org/10.2196/31678)
PMID: [34821561](https://pubmed.ncbi.nlm.nih.gov/34821561/)

©Abraham B Beckers, Johanna T W Snijkers, Zsa Zsa R M Weerts, Lisa Vork, Tim Klaassen, Fabienne G M Smeets, Ad A M Masclee, Daniel Keszthelyi. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 24.11.2021. 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.

Review

Toward Automated Data Extraction According to Tabular Data Structure: Cross-sectional Pilot Survey of the Comparative Clinical Literature

Karl Holub¹, BS; Nicole Hardy¹, MS; Kevin Kallmes¹, MA, JD

Nested Knowledge, St. Paul, MN, United States

Corresponding Author:

Kevin Kallmes, MA, JD

Nested Knowledge

1430 Avon Street N

St. Paul, MN, 55117

United States

Phone: 1 507 271 7051

Email: kevinkallmes@supedit.com

Abstract

Background: Systematic reviews depend on time-consuming extraction of data from the PDFs of underlying studies. To date, automation efforts have focused on extracting data from the text, and no approach has yet succeeded in fully automating ingestion of quantitative evidence. However, the majority of relevant data is generally presented in tables, and the tabular structure is more amenable to automated extraction than free text.

Objective: The purpose of this study was to classify the structure and format of descriptive statistics reported in tables in the comparative medical literature.

Methods: We sampled 100 published randomized controlled trials from 2019 based on a search in PubMed; these results were imported to the AutoLit platform. Studies were excluded if they were nonclinical, noncomparative, not in English, protocols, or not available in full text. In AutoLit, tables reporting baseline or outcome data in all studies were characterized based on reporting practices. Measurement context, meaning the structure in which the interventions of interest, patient arm breakdown, measurement time points, and data element descriptions were presented, was classified based on the number of contextual pieces and metadata reported. The statistic formats for reported metrics (specific instances of reporting of data elements) were then classified by location and broken down into reporting strategies for continuous, dichotomous, and categorical metrics.

Results: We included 78 of 100 sampled studies, one of which (1.3%) did not report data elements in tables. The remaining 77 studies reported baseline and outcome data in 174 tables, and 96% (69/72) of these tables broke down reporting by patient arms. Fifteen structures were found for the reporting of measurement context, which were broadly grouped into: 1×1 contexts, where two pieces of context are reported in total (eg, arms in columns, data elements in rows); 2×1 contexts, where two pieces of context are given on row headers (eg, time points in columns, arms nested in data elements on rows); and 1×2 contexts, where two pieces of context are given on column headers. The 1×1 contexts were present in 57% of tables (99/174), compared to 20% (34/174) for 2×1 contexts and 15% (26/174) for 1×2 contexts; the remaining 8% (15/174) used unique/other stratification methods. Statistic formats were reported in the headers or descriptions of 84% (65/74) of studies.

Conclusions: In this cross-sectional pilot review, we found a high density of information in tables, but with major heterogeneity in presentation of measurement context. The highest-density studies reported both baseline and outcome measures in tables, with arm-level breakout, intervention labels, and arm sizes present, and reported both the statistic formats and units. The measurement context formats presented here, broadly classified into three classes that cover 92% (71/78) of studies, form a basis for understanding the frequency of different reporting styles, supporting automated detection of the data format for extraction of metrics.

(*JMIR Form Res* 2021;5(11):e33124) doi:[10.2196/33124](https://doi.org/10.2196/33124)

KEYWORDS

table structure; systematic review; automated data extraction; data reporting conventions; clinical comparative data; data elements; statistic formats

Introduction

Extracting Data for a Systematic Review

Systematic reviews and meta-analyses of high-quality studies are essential for clinical decision-making [1], guidelines [2], and evidence-based adoption and approval of therapies [3]. Quantitative data extraction is an essential task in the systematic review/meta-analysis process, during which researchers gather patient characteristics, interventions, and outcomes of interest in a common format to support summarization and statistical analysis. The current practice for data extraction is manual review of published manuscripts of studies, with subsequent manual entry of data into a spreadsheet or review software [4]. The manual, work-intensive nature of this task contributes to the high cost in time and money of reviewing the clinical literature.

The time investment and costs of systematic reviews/meta-analyses—which can reach 16 months and US \$141,000 [5] in labor to complete a single review—are the major limiting factors in the synthesis of scientific evidence. The task of data extraction from published comparative studies typically demands 20% of the total review and analysis time, and is subject to high accuracy standards [6,7]. This has led to calls for both improved software systems for systematic reviews/meta-analyses and automation of the data extraction process. However, according to a systematic review of systematic review/meta-analysis extraction automation projects, “no unified information extraction framework [has been] tailored to the systematic review process...[automation] techniques have not been fully utilized to fully or even partially automate the data extraction step of systematic review” [8].

Systematic Review Workflow Software Platforms

Despite the fact that automated data extraction for systematic reviews/meta-analyses has yet to be achieved, several web-based software options currently support part or all of the workflow of a review [9], establishing a systematic approach on which automated data extraction can be modeled. We previously developed a workflow software platform (AutoLit, Nested Knowledge, MN) [10] for performing and presenting systematic reviews and meta-analyses. The data extraction functions of AutoLit are user-driven and focused on extracting descriptive statistics. After articles are retrieved and screened, users read the PDFs of study content and feed extracted data directly into a database, which is used to produce a “living” summary and obtain interpretive statistical outputs. This platform has provided a basis for experimentation with the streamlining of data extraction, the end goal of which is automated identification, parsing, and abstraction of summary statistics reported in medical manuscripts.

Automated Data Extraction Efforts

To begin to solve the problem of automated data extraction, it is first essential to understand the format in which input data is available. PDF manuscripts are the de facto publication medium. Within these manuscripts, the key data regarding the patient population/characteristics, interventions of interest, and outcomes are presented in both the text and data tables. Notably,

the majority of previous extraction efforts have focused on textual extraction [8], despite the varied presentation styles and unstructured nature of both contextual information and the data themselves.

Targeting Extraction from Tables

We hypothesized that data tables, as opposed to free text, represent an ideal target for automated extraction based on the following traits: (1) data in tables are densely concentrated; (2) tables are delimited and typically include structure and a standardized set of contexts not found in free text; (3) tables often report statistics not mentioned in the free text (eg, secondary outcomes); and tables consistently report data with higher precision and full information (eg, dispersion measures and sample sizes).

Existing Standards for Tabular Presentation

Journals and medical research bodies have published standards related to how statistics and tables should be presented [11-13]. Common themes include presenting units for continuous data elements, standardization of statistic formatting (eg, mean [SD]), reporting interventions as full names or standardized abbreviations, and reporting sample sizes used in an analysis.

Despite these guidelines, table formatting standards, both in terms of style and content, vary between journals. This heterogeneity has been noted previously, and a software tool targeted toward authors [14], “tableone,” was created to enable harmonized generation of statistical analyses and tables directly from study results. However, tableone and similar tools are not yet widely adopted to a sufficient extent to meaningfully standardize reporting.

Classifying Tabular Reporting Practices

Given this context, automated tabular extraction depends on understanding the variety of table structures—including the types and frequencies of variation from common formats—in medical manuscripts. A recent systematic review showed that although 14 independent automation projects focused on full-text extraction have been published [15], only one project focused on extraction from tables [8]. Furthermore, although the results of this project were promising, achieving high accuracy in machine learning–based extraction, neither this nor any other study to date has surveyed or classified the structure of tables presented in the clinical literature. Therefore, in this study, we focused on identifying the characteristics that are essential for the automation of extraction of descriptive statistics in tables. This can provide concrete structural characteristics to enable assessment of the generalizability of tabular presentation formats and support the future automation of extraction.

Methods

Sampling Published Comparative Clinical Studies

A cross-sectional sample of clinical study publication records was generated. In brief, published studies tagged as randomized controlled trials (RCTs), as indexed in PubMed, from 2019 were searched using the following term: “randomized controlled trial” [Publication Type] AND 2019/01/01:2020/01/01[dp]. Search results were exported from PubMed on August 9, 2021, using

the Entrez application programming interface [16]; imported to the AutoLit platform; and search results were randomly sampled by index, without replacement, using the R function “sample” to select 100 records. Of these, all published articles were included except those that met the following exclusion criteria: not clinical, not comparative (ie, the publication does

not compare outcomes between patient groups), not in English, protocol only, or no full text available.

Abstracting Tables

Within this sample, table attributes were identified, classified, and summarized. The concepts for data extraction used are summarized in [Table 1](#).

Table 1. Defining tabular concepts.

Concept	Definition	Example
Data element	A characteristic or quality being measured	Mortality
Metric	A measured instance of a data element, a descriptive statistic	2/59 (3.4%)
Arm	A subset of an experiment’s participants that are assigned a specific intervention	Placebo group
Time point	The point(s) in time in an experiment when measurement of data elements is performed	6-month follow-up
Measurement context	The combination of a data element, arm, and time point in experimental reporting	Mortality in the placebo group at 6 months

Data extraction necessarily has to address the attributes of measurement context to discern the meaning of a given metric. Thus, for each patient arm, the intervention and population size must be identified, and for each data element, the unit of measurement and time points must be extracted. Furthermore, metrics must be parsed into their constituent statistics, including (1) continuous metrics as the measure of central tendency and the dispersion measure; (2) dichotomous metrics, namely the subset, total population, and percentage (n/N, %); and (3) categorical metrics, namely the subset and total population.

Tables and their descriptions typically contain at least partial representations of metrics and their measurement contexts. Context can be assigned to dimensions (rows and columns) of the table. For example, [Figure 1](#) displays a “2×1” context, meaning that the rows of the table correspond to 2 nested pieces of context (arms nested in data elements) and the columns correspond to 1 piece of context (time points) [17]. For comparison, [Figure 2](#) displays a “1×1” context, wherein the data elements are labeled on the rows (with corresponding statistical formats) and arms are defined in the columns (time points not presented) [18].

Figure 1. Example of a 2x1 context with arms (blue) nested in data elements (red) on rows and time points (green) on the columns. Table from Chellappa et al [17]. ANOVA: analysis of variance.

TABLE 2: Summary of statistical results of the mixed ANOVA and pooled means and standard deviations of all variable measures.

Variables	Group 1 (TENS-PNF) Group 2 (PNF) Control group (CG)	Mean (standard deviation)			Mixed ANOVA (p value)		
		Pre (1)	3 rd week (2)	5 th week (3)	Interaction (group and time) effect η_p^2 (p value)	Group (G) effect η_p^2 (p value)	Time (T) effect η_p^2 (p value)
Pain VAS (cm)	TENS-PNF	2.7 ± 1.0	0.5 ± 0.6	0.4 ± 0.5	0.608 (0.01)*	0.192 (0.002)*	0.674 (0.01)*
	PNF	2.3 ± 1.0	1.6 ± 0.8	1.6 ± 0.8			
	CG	2.2 ± 0.9	2.0 ± 0.8	2.0 ± 0.8			
Knee to wall (cm)	TENS-PNF	7.5 ± 1.8	10.1 ± 1.2	10.2 ± 1.2	0.661 (0.01)*	0.482 (0.01)*	0.661 (0.01)*
	PNF	6.7 ± 0.8	7.5 ± 0.9	7.2 ± 1.0			
	CG	7.0 ± 0.8	7.1 ± 0.8	7.1 ± 0.8			
Dorsi proprioception	TENS-PNF	2.1 ± 1.2	0.5 ± 0.5	0.4 ± 0.5	0.438 (0.01)*	0.141 (0.01)*	0.390 (0.01)*
	PNF	1.8 ± 1.0	1.5 ± 1.0	1.5 ± 1.0			
	CG	1.9 ± 1.0	1.8 ± 1.0	1.8 ± 1.0			
Plantar proprioception	EG 1	2.5 ± 1.7	0.2 ± 0.5	0.2 ± 0.5	0.520 (0.01)*	0.244 (0.01)*	0.439 (0.01)*
	PNF	2.2 ± 1.9	1.9 ± 1.0	1.9 ± 1.1			
	CG	2.4 ± 1.2	2.4 ± 1.3	2.4 ± 1.2			
Dorsiflexion ROM	TENS-PNF	15.9 ± 1.5	18.6 ± 1.8	18.7 ± 1.8	0.655 (0.01)*	0.064 (0.15) [#]	0.622 (0.01)*
	PNF	16.4 ± 1.8	17.2 ± 1.9	18.7 ± 1.8			
	CG	16.7 ± 1.7	16.8 ± 1.9	16.8 ± 1.8			
Plantarflexion ROM	TENS-PNF	36.4 ± 3.9	43.5 ± 3.7	43.6 ± 3.9	0.729 (0.01)*	0.148 (0.01)*	0.693 (0.01)*
	PNF	38.5 ± 3.3	39.9 ± 3.4	39.3 ± 3.0			
	CG	37.3 ± 4.0	37.5 ± 4.1	37.5 ± 4.3			
Dorsiflexion MS	TENS-PNF	11.1 ± 1.8	13.9 ± 1.7	13.9 ± 1.7	0.741 (0.01)*	0.298 (0.01)*	0.736 (0.01)*
	PNF	11.2 ± 0.8	11.7 ± 0.7	11.6 ± 0.6			
	CG	11.1 ± 0.8	11.3 ± 0.8	11.3 ± 0.8			
Plantarflexion MS	TENS-PNF	13.2 ± 1.7	17.1 ± 1.7	17.1 ± 1.7	0.594 (0.01)*	0.218 (0.01)*	0.541 (0.01)*
	PNF	14.5 ± 0.9	15.1 ± 0.9	14.9 ± 1.0			
	CG	14.1 ± 1.4	14.3 ± 1.5	14.2 ± 1.5			
FADI score	TENS-PNF	94.3 ± 2.5	98.4 ± 1.9	98.6 ± 1.7	0.678 (0.01)*	0.078 (0.098) [#]	0.651 (0.01)*
	PNF	4.7 ± 3.5	5.3 ± 3.4	95.3 ± 3.4			
	CG	5.3 ± 3.8	5.5 ± 3.7	95.5 ± 3.7			

VAS: Visual Analogue Scale; ROM: range of motion; MS: muscle strength; FADI: the Foot and Ankle Disability Index score. *Significant effect ($p \leq 0.05$).
[#]Nonsignificant.

Figure 2. A 1x1 baseline table reporting data elements on rows and arms on columns. Arm sizes are embedded in intervention headers (red), category labels are reported in the data element array indented (blue), and statistic formats are reported in headers (green). Table from Gauto Benitez et al [18]. HFNC: high flow nasal cannula.

Variables		HFNC n (32)	Control n (33)	p
Age	median (range)	5 (2-14)	4 (2-14)	0.537*
Sex	n (%)	22 (68.8)	17 (51.5)	0.156**
		10 (31.3)	16 (48.5)	
Previous crisis	n (%)	30 (93.8)	33 (100)	1****
Maintenance treatment	n (%)	13 (40.6)	15 (45.4)	0.694**
Z Score	mean (SD)	0.45 ± 1.12	0.70 ± 1.53	0.457***
Respiratory rate at admission	mean (SD)	43.8 ± 10.56	46.3 ± 11.62	0.515***
Saturation of O ₂ at admission	median (range)	91 (82-99)	92 (80-98)	0.659*

Table Classification

Classification System

Each included record underwent full text review and was tagged in AutoLit based on the attributes of its (1) table structure, (2) measurement context, and (3) metric information. All tables within a published article were considered for classification via the tagging hierarchy. Tags were assigned on a per-table basis

within each record such that the tags described the attributes of its tables.

Table Structure Reporting

As defined in [Table 2](#), table attributes covered the reporting of baseline and outcome data and table orientation/pagination. If a table did not report any descriptive statistics concerning patient characteristics or outcomes, it was accordingly not tagged.

Table 2. Tagging hierarchy of table structural attributes.

Tags	Applied when
Baseline or demographic characteristics of one or more study arms reported	
In table	The article reports baseline characteristics for the study population in a table, broken out by arm
Participant level	The article reports baseline characteristics for the study population in a table, reported for each participant
No arm-level breakout	The article reports baseline characteristics for the entire study population in a table, with no breakout
Experimental outcomes reported	
In table	The article reports outcomes in a table, including the primary outcome(s), at one or more follow-up time points
Participant level	The article reports outcomes in a table, including the primary outcome(s), at one or more follow-up time points, reported for each participant
Secondary only, in table	The article reports outcomes in a table, but not all the primary outcomes of the study, instead focusing on data points other than primary outcomes
Other table features	
Rotated 90 degrees	One or more baseline or outcome tables in the article is rotated 90 degrees in either direction on the page but is otherwise normal
Multipage	One or more baseline or outcome tables in the article overflows beyond its starting page but is otherwise normal

Measurement Context Reporting

Measurement context tags were applied for all tables for which patient baseline characteristics or outcomes were reported. Contextual information of interest included the pieces of context

per array (where an array represents either a row or column), including data element headers, arm names and arm size reporting, labeling of interventions, and labeling of time points ([Table 3](#)).

Table 3. Tagging hierarchy of measurement context attributes.

Tags	Applied when
Alignment of table dimensions and measurement context^a	
1×1	Only two pieces of context are shown in the table dimensions: one on rows and one on columns (eg, data elements on rows and arms columns)
2×1	Three pieces of context are shown in the table dimensions: two on rows and one on columns (eg, arms nested in data elements on rows, time points on columns)
1×2	Three pieces of context are shown in the table dimensions: one on rows and two on columns
Number of participants in individual arms of the study reported	
Embedded in arm	Arm sizes are reported as part of the arm or intervention label (eg, “Placebo [n=25]”)
Separate array	Arm sizes are reported in a distinct column or row in the table
Header labels corresponding to the intervention(s) applied in each arm of the study	
Full name	The entire name of the intervention(s) for the arm is shown in the arm header
Acronym or abbreviation	An acronym or shortened version of the invention name(s) for the arm is shown in the arm header
Control/experimental	The arm header is labeled with “Control” and “Experimental” or “Treatment” or “Intervention”
Alternate labels	Any header labeling scheme not identified above is used
Header labels corresponding to the time at which the reported data were collected	
Contains unit of time	The time point header contains an amount of time, including units
Pre/post	The time point header is labeled “Pre/Post,” “Before/After,” “Baseline/Follow-up”
Incremental numbered	Time point headers are labeled with numbers or letters in order of time (eg, “t ₁ ,” “t ₂ ”)

^aContext is tagged as “Embedded” when individual header cells include 2 elements of context (eg, “Baseline BMI”).

Metric Reporting

Unlike table structure and measurement context, metric tags were applied per article rather than per table. All baseline and

outcome tables in the article were considered for metric classification (Table 4).

Table 4. Tagging hierarchy of metric attributes.

Tags	Applied when
The format of statistics reported in a metric is displayed^a	
In header	The statistic format or just constituents are reported in the header of the array of metrics
In description or footnotes	The statistic format or constituents are reported in the description or footnotes; these may apply to the entire table or be annotations for arrays
Units of continuous data elements reported^b	
In header	The units of data elements are reported in each array header
In descriptions or footnotes	The units of continuous data elements are reported in the description or footnotes; these may apply to the entire table or be annotations for arrays
Not relevant	The article includes no continuous data elements or the continuous data elements are unitless (eg, scale data)
Pattern defining how the constituent statistics in array of metrics are formatted^c	
Continuous	The format is used for continuous data elements
Dichotomous	The format is used for dichotomous data elements, specifically when only a single category is implied (eg, "Mortality" or "Gender Male")
Categorical	The format is used for categorical data elements; this also applies when a dichotomous data element explicitly lists all categories (eg, "Smoking" has separate arrays for "Yes" and "No")
Method of reporting category labels for categorical data elements^d	
Separate array	Category labels are in an entirely separate (delimited) array from the data element header array
Same array	Category labels are in the data element header array, with no distinction from other data element labels
Same array indented	Category labels are in the data element header array, but are nested under the categorical data element header via white space or list indentation
In cell	Categories are all reported in the same cell (eg, "Gender M/F" with metrics "11/9")

^aIf multiple cases apply, the lowest in the table is the classification.

^bIf multiple cases apply, the lowest in the table is the classification. If units are missing on one or more data elements, this classification should be left empty.

^cThe formats under each tag are created as they are encountered in articles.

^dThis classification is left empty in the event that no categorical data elements are reported.

Statistical Analysis

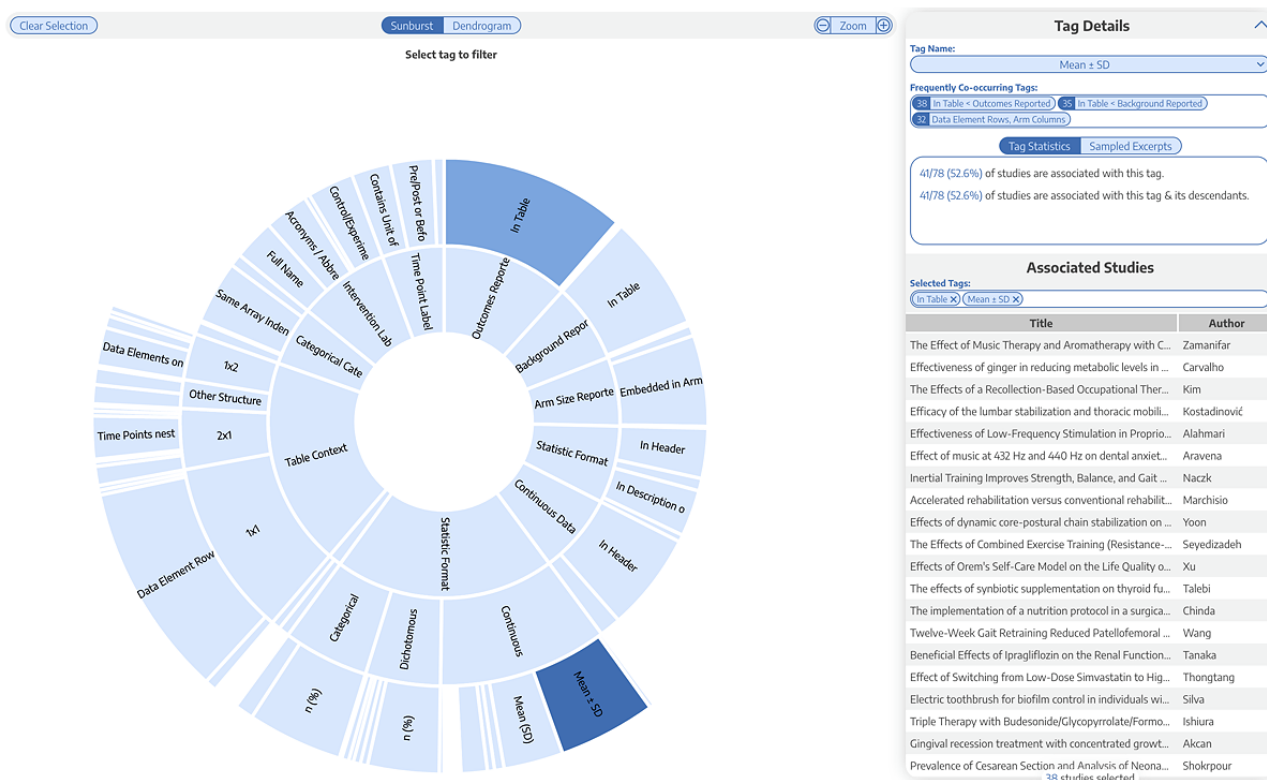
As a pilot study, no power analysis was performed to identify an appropriate sample size. Sample size was estimated to restrict 95% CIs on proportions to $\pm 15\%$ in width. Frequencies were compiled with Boolean queries on tags in AutoLit's Study Inspector. Boolean queries were run on tagged articles using the open source software "btriev" [19] in NodeJS. The counts of results were then compiled and proportion CIs were generated using a normal approximation with the "prop.test" function in R. Inferential statistics on proportions were built with a normal approximation and computed in the R programming language. CIs are reported at the 95% level.

Results

Characteristics of Sampled Articles

Of the sampled 100 records, 12 published articles were excluded for lack of PDF full text, 7 presented nonclinical findings, and 3 were not available in English, leaving 78 articles that were included in this pilot study. A single clinical study was reported in all (78/78, 100%) published articles. Articles were published in 65 distinct journals, with the most frequent journal, *PLoS One*, publishing 9 of the articles. An interactive visualization of all articles tagged using the hierarchical paradigm described above is available on the Nested Knowledge Synthesis page (see Figure 3) [20].

Figure 3. A screenshot of the interactive tagging hierarchy applied across the 78 studies included in this pilot survey. Two filters were applied: “Outcomes Reported In Table” was selected first, and then “Mean ± SD,” meaning the sunburst plot is filtered to studies for which both tags are present. The right menu displays the 38 studies for which this is true as well as statistics about how common the tags in question are across all included studies.



Reporting of Baseline and Outcome Measures in Tables

Baseline and outcome data were reported in 71 of 78 (92%) articles (95% CI 84%-96%). Both baseline and outcome data were presented in tables in 66 of 78 articles (85%, 95% CI 75%-91%), and 77 (99%, 95% CI 93%-100%) articles reported

at least one table of baseline or outcome measures. Standard reporting (with arm-level breakout) was present in 97% (64/66, 95% CI 90%-99%) of tables reporting baseline characteristics and in 96% (69/72, 95% CI 92%-99%) of tables reporting outcomes (Table 5).

Table 5. Baseline and outcome reporting per article and per table.

Type	Frequency per article (N=78), n (%)	Frequency per table (N=174), n (%)
Baseline reported in table	66 (85)	67 (38.5)
Arm-level breakout	64 (97)	65 (97)
No arm-level breakout	2 (3)	2 (3)
Outcomes reported in table	72 (92)	107 (61.5)
Arm-level breakout	69 (96)	104 (97.2)
Secondary only	2 (3)	2 (1.9)
Participant level	1 (1)	1 (0.9)

Among the 174 tables that were found to report either baseline or outcome descriptive statistics, 6 (3.4%, 95% CI 1.6%-7.3%) were rotated 90 degrees and 5 (2.9%, 95% CI 1.2%-6.5%) were multipage.

Reporting of Measurement Context in Tables

Table 6 shows the frequency of measurement contexts using the number of articles reporting one or more baseline or

outcomes tables as the respective denominators. Overall, 48 (62%, 95% CI 51%-72%) articles labeled arms by the full intervention name or an abbreviation of it. Additionally, 22 of the 52 (42%, 95% CI 30%-56%) articles reporting the time point context in tables labeled time points according to the amount of time.

Table 6. Measurement context reporting per article (N=77).

Type	Frequency per relevant ^a article, n (%)
Arm size reported	57 (74)
Embedded in arm	50 (88)
In separate array	6 (11)
In description	1 (1)
Intervention labels reported	77 (100)
Control/experimental	26 (34)
Acronyms or abbreviated	25 (33)
Full name	23 (30)
Alternate labels	3 (4)
Time point labels reported	52 (68)
Pre/post	24 (46)
Contains unit of time	22 (42)
Incremental numbering	6 (12)

^aOne article reported no baseline or outcome data in tables and was thus left out from the measurement context analysis.

In terms of dimensions, the 1×1 context of data elements and arms was most commonly used to report findings, with time points included in only 4 of 99 (4%, 95% CI 1.6%-9.9%) of these tables. Across all context types, arms were most frequently

reported on columns (134/174, 77.0%), data elements on rows (144/174, 82.8%), and time points on columns (38/174, 21.8%) (see [Table 7](#)).

Table 7. Dimensions of tabular reporting of measurement context (N=174).

Context dimensions	Frequency per table, n (%)
1×1	99 (56.9)
DEs ^a on rows, arms on columns	90 (91)
Arms on rows, DEs on columns	5 (5)
Arm on rows, TP ^b on columns	2 (2)
TPs on rows, arms on columns	2 (2)
2×1	34 (19.5)
TPs nested in DEs on rows, arms on columns	18 (53)
Arms nested in DEs on rows, TPs on columns	8 (24)
DEs nested in arms on rows, TPs on columns	2 (6)
Arms nested in TPs on rows, DEs on columns	2 (6)
DEs and TPs embedded in rows, arms on columns	2 (6)
TPs nested in arms on rows, DEs on columns	1 (3)
DEs nested in TPs on rows, arms on columns	1 (3)
1×2	26 (14.9)
DEs on rows, TPs nested in arms on columns	16 (62)
DEs on rows, arms nested in TPs on columns	5 (19)
Arms on rows, TPs nested in DEs on columns	3 (12)
DEs on rows, arms and TPs embedded on columns	2 (8)
Other structure	15 (8.6)
Stratified reporting	8 (53)
Only reports comparative statistics	7 (47)

^aDE: data element.

^bTP: time point.

Reporting of Metrics in Tables

Overall, 75 (97%) articles reported one or more continuous metrics in tables. Fourteen total formats were observed. Mean central tendencies were most frequently reported as mean \pm SD (59%) and mean (SD) (36%), whereas medians were most commonly reported as median (25th percentile-75th percentile)

(52%; see [Table 8](#)). Seven formats were used for means and 7 for medians.

Among 75 articles reporting at least one continuous metric, 64 (85%, 95% CI 76%-92%) contained data elements where units were relevant; of these, 55 (86%, 95% CI 75%-92%) reported units of measurement in table headers.

Table 8. Continuous metric reporting format in tables (N=77).

Continuous metrics reported	Frequency per relevant ^a article, n (%)
Mean	69 (90)
Mean ± SD	41 (59)
Mean (SD)	25 (36)
Mean and SD in separate arrays	2 (3)
Mean	2 (3%)
Mean SD	1 (1%)
Mean (CI; lower-higher)	1 (1%)
Mean ± SD (range; min-max)	1 (1%)
Median	21 (27)
Median (IQR; 25th percentile-75th percentile)	11 (52)
Median (range, min-max)	4 (19)
Median [IQR; 25th percentile-75th percentile]	3 (14)
Median [IQR; 25th percentile, 75th percentile]	1 (5)
Median (IQR; 25th percentile to 75th percentile)	1 (5)
Median (IQR)	1 (5)
Min-Max (median)	1 (5)

^aOne article reported no baseline or outcome data in tables and was thus left out from continuous data characterization.

Regarding dichotomous and categorical metrics, 8 different formats were encountered across 62 articles (Table 9); the format n (%) was most commonly observed for both metric types.

Statistical formats (such as those shown in Table 9) were also commonly represented in tables. In 65 of 77 studies (84%), the generalizable format was given. In 35 articles, the format was presented in the header, whereas in 30 articles, it was presented in the footnote or tabular description.

Table 9. Dichotomous and categorical formats and labels in tables (N=77).

Statistics reported	Frequency per relevant ^a articles, n (%)
Dichotomous	40 (52)
n (%)	30 (75)
n	3 (8)
n/(N-n)	3 (8)
%	3 (8)
n, %	2 (5)
n and % in separate arrays	1 (3)
Categorical	47 (61)
n (%)	40 (85)
n	8 (17)
Categorical labels	47 (61)
Same array, indented	35 (74)
Separate array	7 (15)
In cell	7 (15)
Same array, unindented	1 (2)

^aOne article reported no baseline or outcome data in tables and was thus left out from dichotomous/categorical data characterization.

Among the 47 articles reporting categorical metrics, category label indentation under the data element header was observed

in 35 (74%, 95% CI 60%-85%) articles. Two articles used two distinct methods of category labeling.

Across all data types, statistic formats were reported in tables or their descriptions in 84% (95% CI 75%-91%) of articles.

Characteristics of Articles With High Information Density

Notwithstanding the variety of measurement contexts (with 15 independent formats detected) and statistic formats presented, high information density was common in our dataset. Articles that were found to present maximal information in tables have the following classifications: baseline or outcomes reported in table; arm size reported; intervention labels should be full name or acronyms or abbreviations; when reporting relevant metrics, units of measurement are reported; and statistic format reported.

Thirty-one of the 78 sampled articles (40%, 95% CI 30%-51%) matched these classifications. The most impactful constraint among these classifications was “intervention labels”; if this classification is dropped, 48 (62%, 95% CI 50%-75%) articles matched the maximal-density, most-common formats listed above.

Discussion

Principal Results

In this pilot survey, we found that 85% (66/78) of articles reported both baseline and outcome data in tables, and 99% (77/78) of articles have at least one table of baseline or outcome data. Arm-level (intervention-specific) data were presented in 96% (69/72) of these tables, but there was major heterogeneity in the methods of reporting population size and intervention group names. Tabular dimensions ranged widely, with 15 independent dimensional structures used to report measurement context across 174 tables. Although 1×1 contexts represented the majority of tables, our results suggest that automated context detection will need to contend with a diversity of arrangements.

Similarly, although continuous data were very commonly reported (90% of articles, $n=69/77$) and dichotomous and categorical data were consistently reported (52% and 61% of articles, $n=40/77$ and $n=47/77$ respectively), statistic formats were heterogeneous, with seven formats for means, seven for medians, and six for dichotomous data. Despite the heterogeneity of format, tables provided a consistent, high-density source for baseline and outcome data, and the contexts and formats defined here can be used to refine the expected data presentation for tabular data extraction. We plan to expand upon this pilot study in tabular structure with a review-and-tabular-extraction study, wherein the framework outlined here will be used to classify and extract from underlying articles and the accuracy of extracted metrics will be determined by comparison to manual extraction.

Automated Recognition of Context and Format

The context presented in tables showed the most disappointing rate of reporting information of interest. Although arm sizes were reported in tables in 74% (57/77) of publications, arm interventions and measurement time points were reported in a self-contained manner 60% (46/77) and 40% (31/77) of the time, respectively. If tables are extracted in a completely self-contained manner, with no access to the publication's full

text, we expect only 40% of publications will contain sufficient information in tables alone to complete extraction. Extraction automations will therefore have to receive human input or consult the free text to supplement arm and time point contexts from the tables.

Statistic reporting was heterogeneous in format but extremely commonly reported: statistic format was explicitly reported in tables in up to 90% (69/77) of publications. Even where absent, the format may also be inferred from the metric arrays themselves. Inferred formats may be useful when formats are not reported or as a validation on the detected format. Categorical metrics may produce the most complexity for automations, as category labels are often not distinguished from data element labels by more than whitespace indentation.

Given the commonality of n (%) reporting for dichotomous and categorical data elements, it may be possible to arithmetically derive missing arm sizes from metrics. If not essential, mining the full text may still provide value in validations or completing data. For example, interventions were reported as abbreviations or acronyms in around 30% (23/77) of publications; pattern matching on the abstract or introduction could generate full strings for these shortened versions.

A nontrivial number of publications, around 5% (4/77), contained rotated or multipage tables. Automations should consider tools to identify and apply corrective measures in these cases. Eight percent of tables reported only stratified data or only comparative statistics; although these cases are typically mathematically correctable to arm-level data that meta-analysts desire, automation of these procedures would add complexity. Lastly, although no publication included in our pilot study had missing data, potential missing data must be addressed in any automated workflow: we suggest that for any table where data are missing, the table should be visually presented to users for confirmation.

Previous Research on Tabular Extraction

To date, the scientific literature does not seem to contain any studies giving an overview of the table structure, context presentation, and statistic formats, making this pilot study the first of its kind. The Cochrane Collaboration has created a test data set for automated extraction that may be used to test the accuracy of novel extraction algorithms; however, their data set did not classify tabular structure, instead focusing on providing the test/training data sets and preliminary testing of their own semiautomated extraction system [21].

However, previous authors have proceeded beyond classification and provided approaches for automating tabular extraction. Unlike the approaches reviewed by Jonnalagadda et al [8] and Marshall et al [22], which focus on simplified content-extraction tasks from free text (such as abstracts), Milosevic et al [23] actually tested a preliminary algorithm for tabular extraction. Although this study did not include an overview of the context or statistic formats, the authors achieved an F-score (accuracy) of 82%-92% in the content extraction from a simplified set of HTML tables with 1×1 context formats. The seven-step process for detecting, processing, tagging, and extracting from tables used by Milosevic et al [23] is the most complete tabular

extraction process published to date. The only competing approach focusing on tables in medical publications was from Xu et al [24], who were able to extract drug side effect relationships with 52% accuracy using a statistical classifier. Other than Milosevic et al's [23] pilot study, despite at least 26 approaches attempted in textual extraction [8], automated extraction remains an unmet need for which tabular extraction is a promising and underexamined methodology.

Limitations

We believe our findings will generalize to modern clinical publications owing to the simplicity of our search and applicability of classifications. However, this survey and the AutoLit data extraction framework are applicable only to clinical research publications. Since our search was limited to RCTs, some study types such as observational studies may show different characteristics. Similarly, we did not stratify our results by journal, impact factor, or other factors apart from filtering to RCTs, although journal-related characteristics may affect how representative this pilot is of medical publishing generally.

Additionally, specific fields of research may show characteristics that do not align well with the averaging-across-fields approach used in our study. As a pilot survey, our study did not involve

a power analysis; however, this pilot study can be used to determine sample sizes quantitatively in future research. Lastly, our breakout of contexts and formats is always subject to expansion not covered in this sample, and automations built on the expectation of a limited set of contexts or formats may fail when new presentations of this information are encountered. The test of this framework will be the accuracy of extraction algorithms that employ it compared against existing extraction methods.

Conclusions

In this pilot survey, we found a high density of information in tables, with over 85% (66/78) of articles reporting both background and outcome measures in tables, but with major heterogeneity in presentation of measurement context. Measurement context was most often presented in a 1×1 format, but 15 independent formats were found. Similarly, means and medians were each found in seven independent formats, and dichotomous variables in six. Despite this, high-quality contextual information (intervention labels, arm sizes, units, and statistic formats) were presented in 40% (31/77) of articles. The range of context and statistic formats surveyed here can provide a baseline for future tabular extraction efforts.

Acknowledgments

The authors acknowledge the software development team from Nested Knowledge, Stephen Mead, Jeffrey Johnson, and Darian Lehmann-Plantenberg, for their input in designing the AutoLit workflow. The authors also acknowledge the use of the AutoLit platform and labor funded by Nested Knowledge, Inc, in the completion of this project.

Authors' Contributions

All authors participated in the conception, drafting, and editing of the manuscript. KH was the primary data gatherer and KK reviewed the data gathering for accuracy.

Conflicts of Interest

KH, NH, and KK work for and hold equity in Nested Knowledge, Inc. KK also holds equity in Superior Medical Experts, Inc.

References

1. Gopalakrishnan S, Ganeshkumar P. Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. *J Family Med Prim Care* 2013 Jan;2(1):9-14 [[FREE Full text](#)] [doi: [10.4103/2249-4863.109934](https://doi.org/10.4103/2249-4863.109934)] [Medline: [24479036](https://pubmed.ncbi.nlm.nih.gov/24479036/)]
2. Glasser SP, Duval S. Meta-analysis, evidence-based medicine, and clinical guidelines. In: Glasser SP, editor. *Essentials of clinical research*. Switzerland: Springer, Cham; 2014:203-231.
3. Ciani O, Wilcher B, van Giessen A, Taylor RS. Linking the regulatory and reimbursement processes for medical devices: the need for integrated assessments. *Health Econ* 2017 Feb 31;26(Suppl 1):13-29. [doi: [10.1002/hec.3479](https://doi.org/10.1002/hec.3479)] [Medline: [28139087](https://pubmed.ncbi.nlm.nih.gov/28139087/)]
4. Taylor KS, Mahtani KR, Aronson JK. Summarising good practice guidelines for data extraction for systematic reviews and meta-analysis. *BMJ Evid Based Med* 2021 Jun 25;26(3):88-90. [doi: [10.1136/bmjebm-2020-111651](https://doi.org/10.1136/bmjebm-2020-111651)] [Medline: [33632720](https://pubmed.ncbi.nlm.nih.gov/33632720/)]
5. Michelson M, Reuter K. Corrigendum to "The significant cost of systematic reviews and meta-analyses: A call for greater involvement of machine learning to assess the promise of clinical trials" [*Contemp. Clin. Trials Commun.* 16 (2019) 100443]. *Contemp Clin Trials Commun* 2019 Dec;16:100450 [[FREE Full text](#)] [doi: [10.1016/j.conctc.2019.100450](https://doi.org/10.1016/j.conctc.2019.100450)] [Medline: [31872151](https://pubmed.ncbi.nlm.nih.gov/31872151/)]
6. Haddaway NR, Westgate MJ. Predicting the time needed for environmental systematic reviews and systematic maps. *Conserv Biol* 2019 Apr 24;33(2):434-443. [doi: [10.1111/cobi.13231](https://doi.org/10.1111/cobi.13231)] [Medline: [30285277](https://pubmed.ncbi.nlm.nih.gov/30285277/)]
7. Haddaway N, Westgate M. PredicTER. URL: <https://estech.shinyapps.io/predicter/> [accessed 2021-09-24]

8. Jonnalagadda SR, Goyal P, Huffman MD. Automating data extraction in systematic reviews: a systematic review. *Syst Rev* 2015 Jun 15;4:78 [FREE Full text] [doi: [10.1186/s13643-015-0066-7](https://doi.org/10.1186/s13643-015-0066-7)] [Medline: [26073888](https://pubmed.ncbi.nlm.nih.gov/26073888/)]
9. Van der Mierden S, Tsaïoun K, Bleich A, Leenaars CHC. Software tools for literature screening in systematic reviews in biomedical research. *ALTEX* 2019;36(3):508-517. [doi: [10.14573/altex.1902131](https://doi.org/10.14573/altex.1902131)] [Medline: [31113000](https://pubmed.ncbi.nlm.nih.gov/31113000/)]
10. Nested knowledge. URL: <https://nested-knowledge.com/> [accessed 2021-09-24]
11. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. International Committee of Medical Journal Editors. URL: <http://www.icmje.org/icmje-recommendations.pdf> [accessed 2021-09-24]
12. Arifin WN, Sarimah A, Nors'aadah B, Najib Majdi Y, Siti-Azrin AH, Kamarul Imran M, et al. Reporting statistical results in medical journals. *Malays J Med Sci* 2016 Sep 05;23(5):1-7 [FREE Full text] [doi: [10.21315/mjms2016.23.5.1](https://doi.org/10.21315/mjms2016.23.5.1)] [Medline: [27904419](https://pubmed.ncbi.nlm.nih.gov/27904419/)]
13. Lang TA, Altman DG. Basic statistical reporting for articles published in biomedical journals: the "Statistical Analyses and Methods in the Published Literature" or the SAMPL Guidelines. *Int J Nurs Stud* 2015 Jan;52(1):5-9. [doi: [10.1016/j.ijnurstu.2014.09.006](https://doi.org/10.1016/j.ijnurstu.2014.09.006)] [Medline: [25441757](https://pubmed.ncbi.nlm.nih.gov/25441757/)]
14. Pollard TJ, Johnson AEW, Raffa JD, Mark RG. 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/)]
15. Schmidt L, Olorisade BK, McGuinness LA, Thomas J, Higgins JPT. Data extraction methods for systematic review (semi)automation: a living systematic review. *F1000Res* 2021 May 19;10:401 [FREE Full text] [doi: [10.12688/f1000research.51117.1](https://doi.org/10.12688/f1000research.51117.1)] [Medline: [34408850](https://pubmed.ncbi.nlm.nih.gov/34408850/)]
16. Entrez Programming Utilities Help. National Center for Biotechnology Information. URL: <https://www.ncbi.nlm.nih.gov/books/NBK25501/> [accessed 2021-09-24]
17. Chellappa D, Thirupathy M. Comparative efficacy of low-Level laser and TENS in the symptomatic relief of temporomandibular joint disorders: A randomized clinical trial. *Indian J Dent Res* 2020;31(1):42. [doi: [10.4103/ijdr.ijdr_735_18](https://doi.org/10.4103/ijdr.ijdr_735_18)]
18. Gauto Benítez R, Morilla Sanabria LP, Pavlicich V, Mesquita M. High flow nasal cannula oxygen therapy in patients with asthmatic crisis in the pediatric emergency department. *Rev Chil Pediatr* 2019 Dec;90(6):642-648 [FREE Full text] [doi: [10.32641/rchped.v90i6.1145](https://doi.org/10.32641/rchped.v90i6.1145)] [Medline: [32186587](https://pubmed.ncbi.nlm.nih.gov/32186587/)]
19. btriev. Version 1.0.4. Novel Preposterous Mockery (NPM). 2021. URL: <https://www.npmjs.com/package/btriev/v/1.0.4> [accessed 2021-09-24]
20. Tabular outcomes reporting. Nested Knowledge. URL: <https://nested-knowledge.com/nest/qualitative/443> [accessed 2021-08-23]
21. Norman C, Leeflang M, Névéal A. Data extraction and synthesis in systematic reviews of diagnostic test accuracy: a corpus for automating and evaluating the process. 2018 Presented at: AMIA Annual Symposium Proceedings; November 2018; San Francisco, CA p. 817-826.
22. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst Rev* 2019 Jul 11;8(1):163 [FREE Full text] [doi: [10.1186/s13643-019-1074-9](https://doi.org/10.1186/s13643-019-1074-9)] [Medline: [31296265](https://pubmed.ncbi.nlm.nih.gov/31296265/)]
23. Milosevic N, Gregson C, Hernandez R, Nenadic G. A framework for information extraction from tables in biomedical literature. *Int J Doc Anal* 2019 Feb 15;22(1):55-78. [doi: [10.1007/s10032-019-00317-0](https://doi.org/10.1007/s10032-019-00317-0)]
24. Xu R, Wang Q. Combining automatic table classification and relationship extraction in extracting anticancer drug-side effect pairs from full-text articles. *J Biomed Inform* 2015 Feb;53:128-135 [FREE Full text] [doi: [10.1016/j.jbi.2014.10.002](https://doi.org/10.1016/j.jbi.2014.10.002)] [Medline: [25445920](https://pubmed.ncbi.nlm.nih.gov/25445920/)]

Abbreviations

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 24.08.21; peer-reviewed by MS Aslam, S Pranic; comments to author 17.09.21; revised version received 27.09.21; accepted 07.10.21; published 24.11.21.

Please cite as:

Holub K, Hardy N, Kallmes K

Toward Automated Data Extraction According to Tabular Data Structure: Cross-sectional Pilot Survey of the Comparative Clinical Literature

JMIR Form Res 2021;5(11):e33124

URL: <https://formative.jmir.org/2021/11/e33124>

doi: [10.2196/33124](https://doi.org/10.2196/33124)

PMID: [34821562](https://pubmed.ncbi.nlm.nih.gov/34821562/)

©Karl Holub, Nicole Hardy, Kevin Kallmes. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 24.11.2021. 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

Limited Interaction Targeted Epidemiology of HIV in Sexual and Gender Minority American Adolescents and Adults: Feasibility of the Keeping it LITE Study

Neil Gleason¹, MA; Pedro A Serrano², MPH; Alejandro Muñoz², BA; Audrey French³, PhD; Sybil Hosek^{3,4}, PhD

¹University of Washington, Seattle, WA, United States

²Ruth M Rothstein CORE Center, Cook County Health, Chicago, IL, United States

³Division of Infectious Disease, Stroger Hospital, Chicago, IL, United States

⁴Department of Psychiatry, Stroger Hospital, Chicago, IL, United States

Corresponding Author:

Pedro A Serrano, MPH

Ruth M Rothstein CORE Center

Cook County Health

2020 W Harrison St,

Chicago, IL, 60612

United States

Phone: 1 (312) 572 4500

Email: pserrano@cookcountyhhs.org

Abstract

Background: HIV infection rates among sexual minority men and transgender individuals, particularly adolescents and young adults, remain elevated in the United States despite continued improvement in the HIV public health response. However, there remains a knowledge gap in understanding the barriers faced by this community in receiving HIV care and prevention resources. To address this, the Keeping it LITE study was conducted to assess HIV risk factors and barriers to preventive treatment in a large national cohort of young sexual minority men and transgender individuals at high risk of HIV infection.

Objective: This study aims to evaluate the feasibility of enrolling a large remote cohort, challenges encountered in recruitment, and adjustments made to address these challenges.

Methods: A large national cohort (n=3444) of young sexual minority men and transgender individuals were recruited. Participants were recruited via advertisements on social media; social apps for lesbian, gay, bisexual, transgender, and queer individuals; print advertising; and word-of-mouth. Before enrolling, participants verified their HIV status using an at-home HIV test or by providing their own testing documentation. Descriptive statistics were generated, and a series of logistic regressions were conducted to evaluate demographic differences between recruitment methods, HIV testing methods, and enrollment status.

Results: The Keeping it LITE study was particularly successful in recruiting participants via social media, with over half of the participants recruited from advertisements on social media platforms such as Facebook, Instagram, and Snapchat. Participants were also recruited via word-of-mouth; lesbian, gay, bisexual, transgender, and queer apps (ie, Grindr, Scruff); and print advertisements, and participants recruited from these sources tended to be older and have a higher risk profile. The study was also successful in recruiting a large sample of transgender youth, particularly transgender men and nonbinary individuals. At-home HIV testing was acceptable and more heavily used by younger participants, although several barriers were encountered and overcome in the implementation of this testing. The study had more limited success in recruiting participants aged 13-17 years because of lower enrollment rates and barriers to advertising on social media platforms. The implications of these findings for the future development of HIV research and intervention protocols among sexual minorities and trans youth are discussed.

Conclusions: The methods used in the Keeping it LITE study, particularly recruitment via social media, were found to be feasible and acceptable to participants.

(*JMIR Form Res* 2021;5(11):e30761) doi:[10.2196/30761](https://doi.org/10.2196/30761)

KEYWORDS

social epidemiology; adolescents and young adults; sexual and gender minorities; HIV testing

Introduction

Background

Despite advances in HIV diagnostics, care, and prevention strategies, HIV infection rates among adolescent and young adult sexual minority men and transgender persons continue to rise in the United States. In 2018, 69% of HIV infections in the United States were attributed to male-to-male sexual contact, with sexual minority men ages 25-34 reporting the highest rates of new diagnoses [1]. Although there continues to be a paucity of research on HIV among transgender youth, 2 studies of young transgender women identified self-reported rates of HIV between 19%-22%, and recent epidemiological studies have demonstrated that many trans masculine adults who have sex with cisgender men (transsexual minority men) may have elevated HIV risk [2,3]. In addition, whereas men who have sex with both men and women (eg, bisexual men and straight-identified men who have sex with men) are less likely to be HIV positive than gay men, they are still more than five times as likely to be HIV positive than men who have sex exclusively with women [4]. Although rates of HIV infection have decreased among many risk groups over the past few years, HIV infection continues to disproportionately impact Black and Hispanic or Latinx sexual minority men and transgender persons, and the most striking racial disparities are found among the youngest sexual minority men [1].

Adolescents and young adults are at an increased risk of HIV infection because of biological, cognitive, psychological, and social changes that occur during this distinct developmental phase [5,6]. Adolescence and emerging adulthood can span the late teens through the late twenties and is characterized by change, instability, and an exploration of love, work, and worldviews [7,8]. Multiple co-occurring vulnerabilities, such as lack of power and resources, exacerbate HIV risk during this time, and the vulnerability of youth who are exploring their sexual orientation or coming to terms with their gender identity can be compounded by stigma and rejection from peers, caregivers, and institutions [9]. These vulnerabilities are exacerbated by coepidemics of poverty and violence as well as the syndemics of mental health issues, drug abuse, and victimization. Despite local and national initiatives, including widespread testing, treatment as prevention, educational programming, and the availability of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP), new HIV infections among young sexual minority men and transgender persons persist. In fact, while the number of HIV infections in sexual minority men and transgender persons between 2014 and 2018 decreased among most age groups, annual increases were seen among those aged 25-34 years [1].

Unfortunately, young and very young sexual minority men and transgender persons are less likely to be aware of and avail themselves of HIV testing and effective prevention technologies because of barriers at multiple levels. A significant knowledge gap exists in understanding the role of these barriers in the ongoing epidemic of HIV acquisition among young sexual

minority men and transgender persons. Therefore, there is an urgent need to articulate the epidemiology of HIV acquisition in this population to target culturally and age-appropriate prevention interventions toward those at the highest risk of infection.

Objectives

To address this knowledge gap, the *Keeping it LITE: Exploring HIV Risk in Vulnerable Youth with Limited Interaction* (UG3 AI 133669 & UH3 AI 133676) study was designed to assess HIV risk factors and barriers to preventive treatment in a large national cohort of youth and young adults at high risk of HIV infection. The purpose of this study is to evaluate the feasibility of enrolling (via web-based recruitment) a large, remote cohort of diverse sexual minority men and transgender persons aged 13-34 years (including 20% ages 13-17 years) at high risk of HIV infection and discuss challenges encountered in recruitment and enrollment, and adjustments made to address these challenges. Specifically, we will (1) examine which recruitment methods were most successful for specific demographics; (2) examine the characteristics of individuals who screened eligible for the cohort but did not enroll to identify barriers to enrollment; and (3) evaluate the feasibility of using HIV self-testing as a study procedure.

Methods

Study Development

Youth Advisory Board

During the study development phase, a youth advisory board (YAB) was organized to provide insight into methods of recruitment and retention for young sexual minority men and transgender persons. The YAB consisted of 18 lesbian, gay, bisexual, transgender, and queer (LGBTQ+) youths of color aged 13-24 years. These individuals were recruited via existing relationships with clinical trial networks and from participants enrolled in Keeping it LITE. YAB members reviewed community engagement, outreach, education, recruitment, retention, data collection, incentive, and linkage to care plans and provided feedback.

Recruitment Materials

Advertising and marketing media for the study were developed in collaboration with graphic designers and the YAB. At each stage of development, drafts of recruitment materials were shown to the YAB to assess whether they were understandable, culturally appropriate, and engaging. Examples of feedback from the YAB include suggestions to emphasize visuals over text, to use colors from the LGBTQ pride and trans pride flags to make messaging pop out, to use the same account name on the study website and social media accounts, and to post frequently on social media. The marketing and branding strategy was carefully designed to include all gender and racial or ethnic identities. [Figure 1](#) shows examples of advertisements developed for this study.

Figure 1. Examples of advertisements developed for this study.

KEEPING IT LITE

**L
I
T
E**

#KeepingitLITE is an online study about the life experiences of the LGBTQ+ community.

keepingit.org/mobile
(708) 683-9095

Eligible participants will receive \$50.
Participants may also request a free at-home rapid HIV test kit.

#KeepingitLITE IS AN ONLINE STUDY ABOUT THE HEALTH OF LGBTQ+ PEOPLE.

SIGN UP AT: bit.ly/stayLite
TEXT US AT: (708) 683-9095

Recruitment Procedures

Recruitment advertisements were distributed over social media platforms and mobile apps viewable on all internet-enabled smart devices with specified regions, demographics, and times. The YAB was consulted about methods of recruitment via LGBTQ-specific social media, and they identified Grindr, Jack'd, Scruff, and Adam4Adam as the apps most commonly used by young sexual minority men or transgender persons.

Advertisements were placed on Facebook, Instagram, Snapchat, Spotify, Twitter, YouTube, Grindr, Hornet, Growlr, and Scruff. Print advertisements were also placed in Chicago-area public transit systems and distributed at a limited number of in-person events in Chicago. Finally, participants were also recruited by word-of-mouth, with current participants in the study referring others to join. Participants were given an incentive (US \$20) for each of their enrolled referrals. Recruitment was originally

limited to the Chicago area but later expanded to cities across the United States.

Barriers in Social Media Recruitment

Before launching recruitment, Facebook and Instagram removed the ability to target advertisements through sexual orientation identity [10]. Advertisement targeting was therefore switched from using the built-in sexual orientation indicator to a mix of key ad words to target LGBTQ+ youth and adults (such as media or consumer goods that were popular among this demographic based on suggestions from YAB). Snapchat also changed its advertising policy to block minors from receiving advertisements for clinical trials or research studies. Thus, this medium could only be used to recruit participants aged ≥ 18 years. Finally, some study advertisements were flagged as inappropriate content for minors on Facebook and Instagram. These advertisements typically featured cisgender, transgender, and nonbinary youth holding hands, hugging, giving each other endearing looks, etc. It was noted that photos featuring individuals in summer scenes (eg, beach, swimwear, and shorts) were most likely to be flagged. These advertisements could continue to be used in posts by the study's social media accounts on Facebook and Instagram but were not able to be used in paid advertising.

Prescreening Questionnaire

Study advertisements directed potential participants to a prescreening questionnaire with further information about the study and a series of questions to assess eligibility. To enroll in the study, participants were required to be located in the United States, identify as a cis- or transgender man or transgender woman, or nonbinary person who has sex with persons assigned male at birth, between the ages of 13 and 34 years, report being HIV negative or receiving an HIV diagnosis in the previous year, and report at least one of the following in the last 6 months: (1) condomless anal sex; (2) sex with an HIV-positive partner; or (3) a bacterial sexually transmitted infection (STI). Participants aged < 18 years were enrolled as allowable in their state of residence (6 states do not allow minors to access HIV or STI testing services independently until the ages of 14-16) [11].

In the course of recruitment, it was found that many potential participants under the age of 18 did not meet eligibility criteria because they did not report sufficient sexual risk (ie, reporting no anal sex, HIV-positive partners, or STIs). To facilitate recruitment of the age group, criteria were adjusted to allow participants under 18 to be eligible if they reported engaging in unprotected oral sex with a person assigned male at birth.

Enrollment Procedures

Potential participants who met the inclusion criteria were directed to an electronic consent form to read and sign. Only one enrollment was allowed per IP address to ensure that individuals were enrolled only once. In addition, participants agreed to disclose their name, date of birth, street address, phone number, and email through a secure, Health Insurance Portability and Accountability Act-compliant system. This information was reviewed by the study staff to ensure that the participant was not previously enrolled before a unique participant ID number was assigned.

Assessments

Once informed consent was obtained, the participants completed the HIV status verification. Participants could choose to have an at-home HIV test kit mailed to them or by securely uploading an HIV health record, such as an HIV test lab report, antiretroviral treatment, or PrEP prescription obtained within the last 3 months. Participants then completed a 30-minute baseline assessment with questions assessing demographics, social determinants of health (health insurance status, use of public assistance, immigration status, employment, housing, education, etc), health care experiences, substance use, mental health (social support, attitudes, stigma, discrimination, depression, posttraumatic stress disorder), sexual risk behavior, and HIV prevention engagement. The baseline assessment also included an optional question asking participants how they heard about the study (word-of-mouth, print advertisement, social media, or LGBTQ+ app). The participants completed follow-up evaluations every 6 months. For each follow-up evaluation, they received an automated email prompting them to complete a questionnaire and repeat their HIV testing or status verification.

Retention Strategies

When participants entered their follow-up window (180 days from the last visit), automated email reminders were sent at 1, 3, 7, 14, and 28 days. Participants were messaged semiannually even if they missed one or more previous interactions. On holidays or special awareness days, such as World AIDS Day and National Youth HIV & AIDS Awareness Day, participants were encouraged to participate in contests posted on social media to win small gifts through lotteries.

Measures

HIV Testing

HIV testing was conducted using oral specimen HIV self-testing kits. This testing method was chosen after consultation with the YAB. The YAB expressed a preference for oral specimens over a self-collected blood specimen, and for the ability to upload results via an encrypted web portal over mailing a test back to the researchers. The study staff purchased large volumes of test kits and shipped them to the participants within a week of their request date. The packages included instructions for self-administration in easy-to-understand language and diagrams as well as a Centers for Disease Control and Prevention self-testing educational pamphlet. Participants were instructed to take a picture of the test results and upload it to the encrypted study survey platform. This platform also included links to video tutorials, demonstrating the self-testing procedure. Several ways to securely reach the staff were prominent in the testing instructions with specific instructions should the test be reactive. Participants who experienced challenges receiving their tests by mail received one-on-one troubleshooting with the study staff. This testing method proved inexpensive, with the total cost of the test, packaging, and postage totaling less than US \$20.

To ensure participant privacy and confidentiality with regard to the testing kits, mail packaging was as discreet as possible, and a fully encrypted, secure, Health Insurance Portability and

Accountability Act-compliant web-based data collection program was used to send results. Participants were informed about privacy measures in place for their test data through the informed consent process, and throughout the testing process, participants were encouraged to express feedback or concerns through confidential communication methods. To ensure the privacy and safety of sexual and gender minority participants, we sought and were granted a waiver of the requirement of parental consent for participation.

Adjunct Surveys

Three adjunct surveys were added to the study to evaluate the impact of relevant current events. This included a brief survey in late November 2019-December 2019 to evaluate the effects of misleading advertisements seeking plaintiffs in a lawsuit against Gilead for Truvada toxicities [12]; a survey in May and September of 2020 and May of 2021 to ascertain how the COVID-19 pandemic was affecting the mental, financial, and sexual health of participants [13]; and a short survey in October of 2020 to better understand youth aged 13-17 years who screened eligible but chose not to enroll. The results of the first two surveys are beyond the scope of this study and have been presented and discussed elsewhere [12,13]. The methods of the latter survey are discussed here to provide insight into the recruitment and retention of minor participants.

In October 2020, 246 participants who had consented to participate between December 2017 and December 2019 were between the ages of 13 and 17 at the time of consent but failed to complete baseline HIV testing (which would have completed study enrollment) were invited to participate in a short survey. The survey included practical questions about the mechanics of receiving an HIV test, privacy concerns, data use, electronic incentives, as well as issues of stigma, disclosure, risk perception, and fear of test results.

Participant Compensation

Participants were compensated US \$50 per study visit after the HIV test result or status verification was received, and the survey

was completed. Participants received compensation through a web-based gift card distribution system that allowed them to choose between a variety of electronic gift cards or a prepaid debit card. The choice to use this web-based system was made after receiving feedback from participants that they wanted to receive their payments as quickly as possible and after consultation with the YAB. For the ad hoc surveys, the study team experimented with random lottery and flat-rate compensation, but ultimately found no difference in response rates. Participants were also compensated with US \$20 for completing adjunct surveys.

Statistical Analysis

Descriptive statistics were generated for the data, including frequencies, means, SDs, and odds ratios (ORs) for participant demographics and recruitment methods, testing methods, and enrollment status. A series of logistic regressions were conducted to evaluate differences in participant demographics between recruitment methods, HIV testing methods, and enrollment status. Analyses were performed using JASP version 0.14.1 [14].

Results

Sample Characteristics

The demographic characteristics of the enrolled participants are shown in [Table 1](#). Participants were recruited from December 2017 to December 2019 ([Figure 2](#)) and recruited from across the United States ([Figure 3](#)). Most participants identified as cisgender men (2613/3444, 75.87%), with most cisgender men in the study identifying as gay (2089/2613, 97.94%). Furthermore, 1.91% (66/3444) of participants identified as transgender women, 8.04% (277/3444) identified as transgender men, and 14.16% (488/3444) of participants identified as genderqueer or gender nonconforming. Nearly half of the participants were identified as persons of color (1631/3444, 47.35%).

Table 1. Characteristics of enrolled participants (N=3444).

Characteristics	Total (N=3444)	White (n=1813)	Latinx (n=680)	Black (n=417)	Asian or Pacific Islander (n=154)	Other (n=380)
Age (years), mean (SD)	24.6 (4.8)	25.0 (4.8)	24.2 (4.9)	24.4 (4.7)	24.6 (4.8)	23.7 (4.8)
Sexual orientation and gender, n (%)						
Cisgender men						
Gay	2098 (60.91)	1165 (64.25)	416 (61.18)	227 (54.43)	100 (64.94)	190 (50)
Other	506 (14.69)	201 (11.09)	119 (17.5)	88 (21.1)	25 (16.23)	73 (19.21)
Straight	9 (0.26)	4 (0.22)	1 (0.15)	2 (0.48)	0 (0)	2 (0.53)
Nonbinary						
Gay	138 (4)	54 (2.97)	29 (4.26)	34 (8.15)	4 (2.59)	17 (4.47)
Other	345 (10.02)	184 (10.15)	56 (8.23)	43 (10.31)	15 (9.74)	47 (12.37)
Straight	5 (0.14)	0 (0)	1 (0.15)	2 (0.48)	0 (0)	2 (0.53)
Trans men						
Gay	54 (1.57)	36 (1.99)	7 (1.03)	1 (0.24)	2 (1.29)	8 (2.11)
Other	213 (6.18)	129 (7.12)	33 (4.85)	10 (2.39)	8 (5.19)	33 (8.68)
Straight	10 (0.29)	6 (0.33)	3 (0.44)	0 (0)	0 (0)	1 (0.26)
Trans women						
Gay	12 (0.35)	6 (0.33)	2 (0.29)	3 (0.72)	0 (0)	1 (0.26)
Other	44 (1.28)	28 (1.54)	7 (1.03)	5 (1.19)	0 (0)	4 (1.05)
Straight	10 (0.29)	0 (0)	6 (0.88)	2 (0.48)	0 (0)	2 (0.5)
Location^a, n (%)						
EHE2^b						
Midwest	1124 (32.64)	526 (29.01)	237 (34.85)	200 (48.96)	51 (33.12)	110 (28.95)
South	449 (13.04)	226 (12.46)	80 (11.76)	82 (19.66)	13 (8.44)	48 (12.63)
Northeast	183 (5.31)	89 (4.91)	35 (5.15)	18 (4.32)	18 (11.69)	23 (6.05)
West	241 (6.99)	106 (5.85)	76 (11.18)	14 (3.36)	19 (12.34)	26 (6.84)
Non-EHE						
Midwest	572 (16.61)	371 (20.46)	75 (11.03)	35 (8.39)	20 (12.98)	71 (10.44)
South	358 (10.39)	195 (10.76)	73 (10.74)	45 (10.79)	8 (5.19)	37 (9.74)
Northeast	291 (8.45)	174 (9.59)	50 (7.35)	19 (4.56)	12 (7.79)	36 (9.47)
West	226 (6.56)	126 (6.95)	54 (7.94)	4 (0.96)	13 (8.44)	29 (7.63)
HIV status, n (%)						
HIV negative	2561 (74.36)	1388 (76.56)	507 (74.56)	258 (61.87)	117 (75.97)	291 (76.58)
HIV negative and on PrEP ^c or PEP ^d	769 (22.32)	407 (22.45)	151 (22.21)	99 (23.74)	36 (23.38)	76 (20)
HIV positive	114 (3.31)	18 (0.99)	22 (3.23)	60 (14.39)	1 (0.65)	13 (3.42)
Permanent housing^a, n (%)						
Yes	3208 (93.15)	1742 (96.08)	618 (90.88)	347 (83.21)	149 (96.75)	352 (92.63)
No	204 (5.92)	57 (3.14)	55 (8.09)	65 (15.59)	4 (2.59)	23 (6.05)
Education^a, n (%)						
≤Bachelor's degree	1548 (44.95)	899 (49.59)	276 (42.59)	141 (33.81)	95 (61.69)	137 (36.05)
≥Bachelor's degree	1670 (48.49)	789 (43.52)	372 (54.71)	253 (60.67)	50 (32.48)	206 (54.21)

Characteristics	Total (N=3444)	White (n=1813)	Latinx (n=680)	Black (n=417)	Asian or Pacific Islander (n=154)	Other (n=380)
Insurance status^a, n (%)						
Insured	2602 (75.56)	1423 (78.49)	501 (73.68)	278 (66.67)	129 (83.77)	271 (71.32)
Not Insured	334 (9.69)	118 (6.51)	104 (15.29)	66 (15.83)	7 (4.54)	39 (10.26)
Employment^a, n (%)						
Yes	1937 (56.24)	1098 (60.56)	365 (53.68)	216 (51.79)	74 (48.05)	184 (48.42)
No	1026 (29.79)	457 (25.21)	236 (34.71)	155 (37.17)	43 (27.92)	135 (35.52)
Any income^a, n (%)						
Yes	2786 (80.89)	1476 (81.41)	563 (82.79)	322 (77.22)	131 (85.06)	294 (77.37)
No	436 (12.66)	215 (11.86)	86 (12.65)	72 (17.27)	14 (9.09)	49 (12.89)
Number of risk criteria^e, n (%)						
2 or more	1241 (36.03)	584 (32.21)	256 (37.65)	208 (49.88)	51 (33.12)	142 (37.37)
Less than 2	2203 (63.96)	1229 (67.79)	424 (62.35)	209 (50.12)	103 (66.88)	238 (62.63)
PrEP indicated^f, n (%)						
Yes	2972 (86.29)	1594 (87.92)	599 (88.09)	328 (78.66)	138 (89.61)	313 (82.37)
No	472 (13.7)	219 (12.08)	81 (11.91)	89 (21.34)	16 (10.39)	67 (17.63)
Substance use^{a,g}, n (%)						
Yes	533 (15.48)	276 (15.22)	109 (16.03)	49 (11.75)	21 (13.64)	78 (20.53)
No	2133 (61.93)	1206 (66.52)	398 (58.53)	238 (57.07)	90 (58.44)	201 (52.89)

^aDenotes missing data.

^bEHE: ending the HIV epidemic priority jurisdictions [15].

^cPrEP: pre-exposure prophylaxis.

^dPEP: postexposure prophylaxis.

^eRisk criteria: (1) inconsistent condom use, (2) HIV-positive partner, and (3) bacterial sexually transmitted infection in the past 6 months.

^fPrEP indication is based on Centers for Disease Control and Prevention's draft of clinical practice guidelines [16]: HIV negative, recent anal or vaginal sex, and one of the following: inconsistent condom use, HIV-positive partner, or bacterial sexually transmitted infection in past 6 months.

^gSubstance use: cocaine, meth, or heroin use in past 6 months.

Figure 2. Enrollment activity by study quarter.

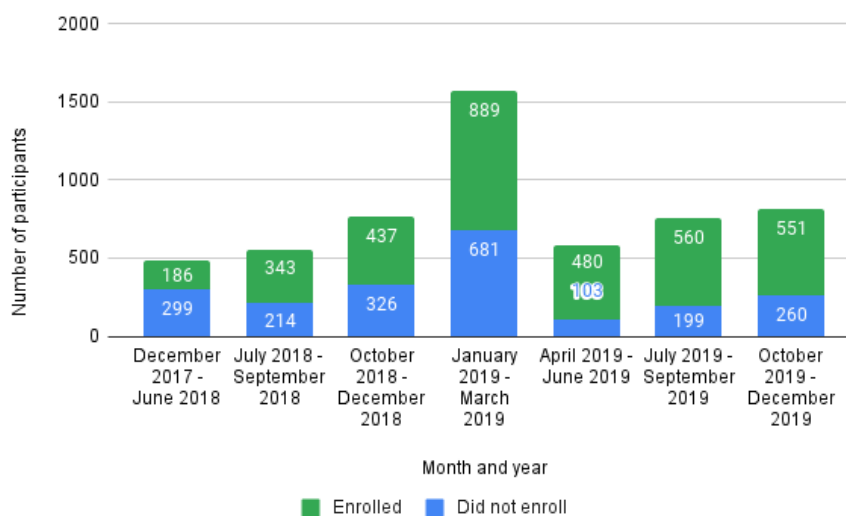
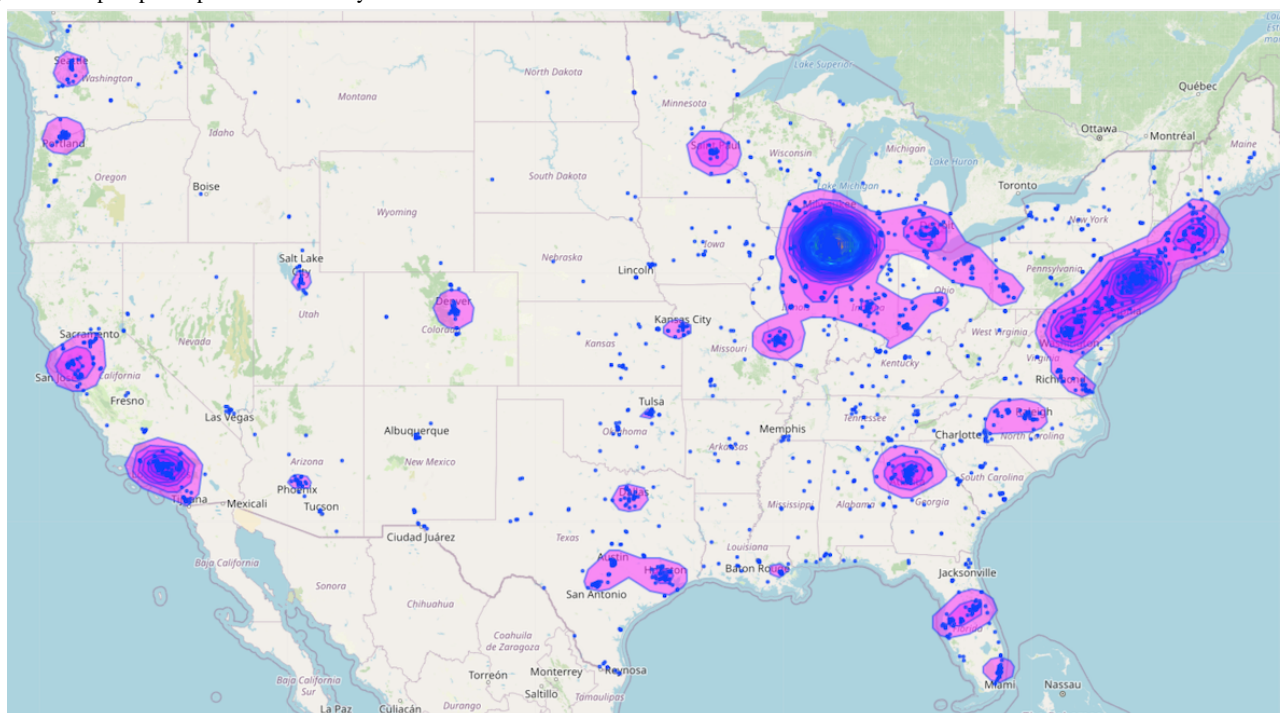


Figure 3. Map of participant enrollment by location.

Social Media Engagement and Recruitment Sources

During the study recruitment period, digital advertisements for the study were seen (ie, displayed on someone's social media feed) over 2 million times and clicked open over 720,000 times. Social media platforms such as Snapchat, Instagram, and Facebook accounted for a majority (about 9 out of 10) of clicks, whereas about 1 in 10 clicks were from dating apps such as Grindr, Hornet, Growlr, and Scruff. The study Facebook page accrued more than 4200 likes, and the study Instagram account accrued more than 2700 followers. The website for the study received over 35,000 unique visitors.

A total of 2649 participants reported how they heard about the study. Descriptive statistics, ORs, and results of the logistic regression analysis are displayed in [Table 2](#). The most

commonly reported method of recruitment was social media (1684/2649, 63.57%), although some methods were more effective for certain demographics than others. A series of logistic regressions indicated advertisements on LGBT+ apps such as Grindr more commonly enrolled older participants with a higher risk profile, whereas advertisements on social media apps such as Instagram and Snapchat tended to attract younger participants with lower risk profiles. Cisgender men were more likely than transgender participants to hear about the study from Grindr or other LGBT apps, and participants aged <20 years were more likely to hear about the study from social media than participants ≥ 20 years. Word-of-mouth and LGBTQ+ app methods also tended to enroll more participants who were HIV positive or using PrEP or PEP. The results of the multivariate model confirmed the independence of these associations ([Table 2](#)).

Table 2. Participant characteristics by recruitment method (n=2649).

Characteristics	Total	Social media	Word-of-mouth	LGBTQ+ ^a apps	Print ads	Social media versus all other, univariate OR ^b (95% CI)	P value ^c	Multivariate OR (95% CI)
Total, n (%)	2649 (100)	1684 (63.61)	499 (18.81)	341 (12.91)	125 (4.61)	— ^d	—	—
Age (years), mean (SD)	24.6 (4.8)	24.3 (4.9)	24.8 (4.6)	26.6 (4.5)	23.5 (4.9)	0.96 (0.94-0.97)	<.001	0.97 (0.95-0.98)
30-34	483 (18.2)	282 (16.7)	85 (17)	97 (28.4)	19 (15.2)	Reference		
25-29	860 (32.5)	506 (30.0)	182 (36.5)	141 (41.3)	31 (24.8)	1.02 (0.81-1.28)		
20-24	858 (32.4)	570 (33.8)	162 (32.5)	77 (22.6)	49 (39.2)	1.41 (1.12-1.78)		
18-19	243 (9.2)	165 (9.8)	42 (8.4)	22 (6.5)	14 (11.2)	1.51 (1.09-2.09)		
13-17	205 (7.7)	161 (9.6)	28 (5.6)	4 (1.2)	12 (9.6)	2.61 (1.78-3.81)		
Race or gender, n (%)								
Sexual minority men							<.001	
White	1091 (41.21)	725 (43.13)	168 (33.68)	154 (45.21)	44 (35.18)	Reference		Reference
Latinx	434 (16.43)	276 (16.41)	71 (14.21)	68 (19.89)	19 (15.21)	0.88 (0.70-1.11)		0.82 (0.63-1.06)
Black	250 (9.37)	122 (7.18)	68 (13.57)	47 (13.78)	13 (10.45)	0.48 (0.36-0.64)		0.58 (0.42-0.80)
Asian	102 (3.89)	69 (4.08)	20 (4.03)	12 (3.54)	1 (0.79)	1.06 (0.68-1.63)		1.06 (0.66-1.69)
Other	206 (7.78)	137 (8.12)	37 (7.42)	23 (6.71)	9 (7.21)	1.00 (0.73-1.37)		1.01 (0.72-1.43)
Trans or nonbinary								
White	297 (11.22)	199 (11.77)	66 (13.23)	15 (4.56)	17 (13.61)	1.03 (0.78-1.35)	—	0.89 (0.66-1.20)
Latinx	101 (3.81)	54 (3.21)	29 (5.79)	9 (2.61)	9 (7.22)	0.58 (0.38-0.87)	—	0.72 (0.45-1.14)
Black	67 (2.45)	35 (2.13)	23 (4.56)	3 (0.87)	6 (4.76)	0.55 (0.34-0.91)	—	0.90 (0.52-1.57)
Asian	20 (0.82)	12 (0.65)	2 (0.36)	6 (1.81)	0 (0)	0.76 (0.31-1.87)	—	0.85 (0.31-2.33)
Other	81 (3.13)	55 (3.33)	15 (3.02)	4 (1.17)	7 (5.61)	1.07 (0.66-1.73)	—	0.85 (0.50-1.45)
Location^e, n (%)								
EHE^f							<.001	
Midwest	735 (27.71)	252 (15.04)	283 (56.74)	96 (28.19)	104 (83.21)	Reference		Reference
South	384 (14.51)	291 (17.34)	43 (8.62)	45 (13.16)	5 (4.02)	6.00 (4.54-7.93)		5.96 (4.48-7.92)
Northeast	150 (5.66)	110 (6.47)	18 (3.61)	19 (5.62)	3 (2.37)	5.27 (3.56-7.81)		5.54 (3.71-8.26)
West	209 (7.92)	154 (9.12)	16 (3.18)	38 (11.11)	1 (0.77)	5.37 (3.81-7.57)		5.19 (3.66-7.37)
Non-EHE								
Midwest	429 (16.21)	295 (17.53)	68 (13.62)	60 (17.61)	6 (4.78)	4.22 (3.27-5.44)	—	3.54 (2.72-4.60)
South	315 (11.87)	245 (14.45)	36 (7.21)	33 (9.72)	1 (0.76)	6.71 (4.94-9.11)	—	6.14 (4.50-8.39)
Northeast	232 (8.81)	185 (11.02)	19 (3.76)	24 (7.02)	4 (3.21)	7.54 (5.29-10.75)	—	6.56 (4.58-9.41)
West	195 (7.41)	152 (9.02)	16 (3.21)	26 (7.63)	1 (0.78)	6.78 (4.67-9.82)	—	6.14 (4.20-8.96)
HIV status, n (%)								
HIV negative	1969 (74.3)	1321 (78.42)	332 (66.45)	221 (64.78)	95 (76.02)	Reference	<.001	Reference
PrEP ^g or PEP ^h	586 (22.1)	320 (19.05)	144 (28.9)	99 (29.0)	23 (18.34)	0.59 (0.49-0.71)		0.78 (0.63-0.97)

Characteristics	Total	Social media	Word-of-mouth	LGBTQ+ ^a apps	Print ads	Social media versus all other, univariate OR ^b (95% CI)	P value ^c	Multivariate OR (95% CI)
HIV positive	94 (3.45)	43 (2.63)	23 (4.6)	21 (6.2)	7 (5.6)	0.41 (0.27-0.63)		0.66 (0.40-1.06)
Number of risk criteriaⁱ, n (%)								
2 or more	989 (37.31)	558 (33.08)	234 (46.89)	160 (46.91)	37 (29.56)	Reference	<.001	Reference
Less than 2	1660 (62.72)	1126 (66.91)	265 (53.13)	181 (53.12)	88 (70.34)	1.63 (1.38-1.92)	—	1.19 (0.98-1.45)
PrEP indicated^j, n (%)								
Yes	2371 (89.45)	1518 (90.1)	440 (88.21)	315 (92.42)	98 (78.42)	Reference	.16	—
No	278 (10.51)	166 (9.88)	59 (11.83)	26 (7.58)	27 (21.57)	0.83 (0.65-1.07)	—	—
Substance use^{e,k}, n (%)								
Yes	416 (20.56)	253 (15.02)	81 (20.71)	62 (24.17)	20 (20.41)	Reference	.32	—
No	1599 (79.34)	1015 (60.27)	311 (79.32)	195 (75.87)	78 (79.58)	1.12 (0.90-1.40)	—	—

^aLGBTQ+: lesbian, gay, bisexual, transgender, and queer.

^bOR: odds ratio.

^cResults of logistic regression (outcome variable=recruited via social media vs recruited via other method).

^dNot available.

^eDenotes missing data.

^fEHE: ending the HIV epidemic priority jurisdictions [15].

^gPrEP: pre-exposure prophylaxis.

^hPEP: postexposure prophylaxis.

ⁱRisk criteria: (1) inconsistent condom use, (2) HIV-positive partner, and/or (3) bacterial sexually transmitted infection in the past 6 months.

^jPre-exposure prophylaxis indication is based on Centers for Disease Control and Prevention's draft of clinical practice guidelines [16]: HIV negative, recent anal or vaginal sex, and one of the following: inconsistent condom use, HIV-positive partner, or bacterial sexually transmitted infection in past 6 months.

^kSubstance use: cocaine, meth, or heroin use in past 6 months ≥ 2 risk criteria: (1) inconsistent condom use, (2) HIV-positive partner, and/or (3) bacterial sexually transmitted infection in the past 6 months.

Qualification and Enrollment

Table 3 displays the enrollment status of all 11,821 participants who completed eligibility screenings. A series of logistic regressions indicated that younger eligible participants were less likely than older eligible participants to enroll in the study. Participants aged 13-17 years were significantly less likely to enroll compared with participants aged 30-34 years (OR 0.44, 95% CI 0.35-0.56). In addition, compared with eligible White cisgender male-identified participants, Asian (OR 0.88, 95%

CI 0.62-1.13), Black (OR 0.70, 95% CI 0.58-0.85), and Latinx (OR 0.75, 95% CI 0.64-0.89) participants were less likely to enroll, and transgender and nonbinary individuals were less likely to enroll than cisgender participants. Eligible HIV-positive participants were 1.67 (95% CI 1.20-2.33) times more likely and HIV negative participants on PrEP or PEP were 15.73 (95% CI 11.3-21.9) times more likely to enroll than HIV negative participants not on PrEP or PEP. The results of the multivariate model confirmed the independence of these associations (Table 3).

Table 3. Characteristics of participants by enrollment status (n=11,821).

Characteristics	Total	Did not qualify	Did not enroll	Enrolled	Enrolled versus did not enroll, univariate OR ^a (95% CI)	P value ^b	Multivariate OR (95% CI)
Total, n (%)	11,821 (100)	6295 (53.25)	2082 (17.62)	3444 (29.13)	— ^c	—	—
Age (years), mean (SD)	25.1 (8.1)	25.9 (10.0)	23.6 (4.9)	24.6 (4.8)	1.04 (1.03-1.06)	<.001	1.02 (1.00-1.03)
>34	778 (6.6)	778 (12.4)	0 (0)	0 (0)	—		
30-34	1593 (13.5)	681 (10.8)	293 (14.1)	619 (18)	Reference		
25-29	2792 (23.6)	1181 (18.8)	546 (26.2)	1065 (30.9)	0.92 (0.78-1.1)		
20-24	4017 (34.0)	2049 (32.5)	777 (37.3)	1191 (34.6)	0.73 (0.61-0.86)		
18-19	1386 (11.7)	827 (13.1)	220 (10.6)	339 (9.8)	0.73 (0.59-0.91)		
13-17	1252 (10.6)	776 (12.3)	246 (11.8)	230 (6.7)	0.44 (0.35-0.56)		
<13	3 (0)	3 (0)	0 (0)	0 (0)	—		
Race or gender, n (%)							
Sexual minority men						<.001	
White	3726 (31.45)	1671 (26.48)	685 (32.88)	1370 (39.78)	Reference		Reference
Latinx	1405 (11.87)	514 (8.18)	355 (17.09)	536 (15.57)	0.75 (0.64-0.89)		0.79 (0.67-0.94)
Black	892 (7.52)	349 (5.48)	226 (10.87)	317 (9.16)	0.70 (0.58-0.85)		0.66 (0.53-0.81)
Asian	322 (2.69)	122 (1.89)	75 (3.59)	125 (3.62)	0.83 (0.62-1.13)		0.85 (0.62-1.17)
Other	735 (6.16)	295 (4.68)	175 (8.36)	265 (7.66)	0.76 (0.61-0.94)		0.79 (0.63-0.99)
Trans or nonbinary							
White	2476 (20.87)	1773 (28.16)	260 (12.49)	443 (12.89)	0.85 (0.71-1.02)	—	1.12 (0.92-1.36)
Latinx	737 (6.19)	473 (7.52)	120 (5.76)	144 (4.19)	0.60 (0.46-0.78)	—	0.71 (0.54-0.94)
Black	824 (7.04)	627 (10.02)	97 (4.67)	100 (2.91)	0.52 (0.38-0.69)	—	0.60 (0.44-0.82)
Asian	147 (1.23)	108 (1.72)	10 (0.49)	29 (0.77)	1.45 (0.70-2.99)	—	1.85 (0.88-3.85)
Other	557 (4.72)	363 (5.77)	79 (3.8)	115 (3.31)	0.73 (0.54-0.98)	—	0.96 (0.70-1.32)
Location^d, n (%)							
EHE^e						.71	—
Midwest	3666 (31.02)	1893 (30.08)	649 (31.32)	1124 (32.6)	Reference		
South	1487 (12.57)	769 (12.21)	269 (13)	449 (13.03)	0.96 (0.81-1.15)		
Northeast	592 (5.02)	300 (4.79)	109 (5.29)	183 (5.31)	0.97 (0.75-1.25)		
West	772 (6.52)	373 (5.89)	158 (7.62)	241 (6.8)	0.88 (0.70-1.10)		
Non-EHE							
Midwest	2431 (20.62)	1486 (23.61)	373 (18.01)	572 (16.56)	0.89 (0.75-1.04)	—	—
South	1121 (9.47)	558 (8.86)	205 (9.91)	358 (10.42)	1.01 (0.83-1.23)	—	—
Northeast	1011 (8.59)	559 (8.88)	161 (7.81)	291 (8.42)	1.04 (0.84-1.29)	—	—
West	730 (6.2)	356 (5.69)	148 (7.12)	226 (6.6)	0.88 (0.70-1.11)	—	—
HIV status, n (%)							
HIV negative, not on PrEP ^f or PEP ^g	10,536 (89.12)	5984 (95.08)	1991 (95.59)	2561 (74.36)	Reference	<.001	Reference
HIV negative; on PrEP or PEP	807 (6.82)	0 (0)	38 (1.77)	769 (22.31)	15.73 (11.3-21.9)	—	16.42 (11.7-23.05)
HIV positive	478 (4.03)	311 (4.92)	53 (2.53)	114 (3.27)	1.67 (1.20-2.33)	—	2.25 (1.57-3.2)

Characteristics	Total	Did not qualify	Did not enroll	Enrolled	Enrolled versus did not enroll, univariate OR ^a (95% CI)	P value ^b	Multivariate OR (95% CI)
Number of risk criteria^h, n (%)							
3	649 (5.48)	148 (2.39)	170 (8.22)	331 (9.61)	Reference	<.001	Reference
2	1868 (15.77)	539 (8.61)	419 (20.07)	910 (26.38)	1.12 (0.90-1.39)		1.64 (1.28-2.11)
1	4905 (41.45)	1876 (29.83)	1186 (57.03)	1843 (53.51)	0.80 (0.65-0.97)		1.55 (1.22-1.96)
0	4399 (37.23)	3732 (59.33)	307 (14.67)	360 (10.46)	0.60 (0.47-0.77)		1.4 (1.05-1.88)

^aOR: odds ratio.

^bResults of logistic regression (outcome variable=enrolled vs did not enroll).

^cNot available.

^dDenotes missing data.

^eEHE: ending the HIV epidemic priority jurisdictions [15].

^fPrEP: pre-exposure prophylaxis.

^gPEP: postexposure prophylaxis.

^hRisk criteria: (1) inconsistent condom use, (2) HIV-positive partner, and/or (3) bacterial sexually transmitted infection in the past 6 months.

HIV Testing Methods

Table 4 displays HIV testing methods used by participants across demographics. A majority of participants (2573/3444, 74.71%) opted to use the at-home HIV testing kits provided by the study. A series of logistic regressions indicated younger participants used the at-home tests at higher rates than older participants, with participants aged 13-17 years being 6.38 (95% CI 6.61-11.25) times more likely to use at-home tests than participants aged 30-34 years. Compared with cisgender White

men, cisgender Black men were 0.46 (95% CI 0.36-0.59) times as likely and cisgender Asian men were 0.62 (95% CI 0.42-0.92) times as likely to use at-home tests. Trans and nonbinary participants were more mixed, with White participants having greater odds (OR 1.54, 95% CI 1.17-2.04) and Black participants having lower odds (OR 0.58, 95% CI 0.38-0.90) of using at-home tests. The results of the multivariate model confirmed the independence of these associations except for age and substance use (Table 4).

Table 4. Characteristics of participants using HIV self-testing versus other HIV status verification (n=3444).

Characteristics	Total	HIV self-test	Other status verification	Univariate OR ^a (95% CI)	P value ^b	Multivariate OR (95% CI)
Total, n (%)	3444 (100)	2573 (74.71)	871 (25.29)	— ^c	—	—
Age (years), mean (SD)	24.6 (4.8)	24.3 (4.9)	25.5 (4.4)	0.95 (0.93-0.96)	<.001	1.02 (0.99-1.04)
30-34	619 (18.0)	438 (17.0)	181 (20.8)	Reference	<.001	
25-29	1065 (30.9)	761 (29.6)	304 (34.9)	1.03 (0.83-1.29)	—	
20-24	1191 (34.6)	883 (34.3)	308 (35.4)	1.18 (0.95-1.47)	—	
18-19	339 (9.8)	275 (10.7)	64 (7.3)	1.78 (1.29-2.45)	—	
13-17	230 (6.7)	216 (8.4)	14 (1.6)	6.38 (3.61-11.25)	—	
Race or gender, n (%)						
Sexual minority men					<.001	
White	1370 (39.78)	1042 (40.49)	328 (37.67)	Reference		Reference
Latinx	536 (15.56)	400 (15.55)	136 (15.61)	0.93 (0.74-1.17)		1.09 (0.80-1.48)
Black	317 (9.2)	188 (7.31)	129 (14.83)	0.46 (0.36-0.59)		0.64 (0.44-0.92)
Asian	125 (3.63)	83 (3.23)	42 (4.82)	0.62 (0.42-0.92)		0.58 (0.35-1.97)
Other	265 (7.69)	197 (7.66)	68 (7.78)	0.91 (0.67-1.23)		1.09 (0.72-1.63)
Trans or nonbinary						
White	443 (12.86)	368 (14.3)	75 (8.56)	1.54 (1.17-2.04)	—	1.00 (0.70-1.43)
Latinx	144 (4.18)	115 (4.47)	29 (3.34)	1.25 (0.82-1.91)	—	1.10 (0.63-1.91)
Black	100 (2.9)	65 (2.53)	35 (4.02)	0.58 (0.38-0.90)	—	0.53 (0.27-1.04)
Asian	29 (0.84)	22 (0.86)	7 (0.78)	0.99 (0.42-2.34)	—	0.42 (0.14-1.26)
Other	115 (3.34)	93 (3.61)	22 (2.45)	1.33 (0.82-2.15)	—	0.92 (0.50-1.69)
Location^d, n (%)						
EHE^e					<.001	
Midwest	1124 (32.63)	801 (31.13)	323 (37.07)	Reference		Reference
South	449 (13.04)	329 (12.79)	120 (13.78)	1.11 (0.86-1.41)		0.72 (0.52-1.00)
Northeast	183 (5.31)	105 (4.08)	78 (9.03)	0.54 (0.39-0.75)		0.47 (0.30-0.71)
West	241 (6.99)	156 (6.06)	85 (9.82)	0.74 (0.55-0.99)		0.44 (0.30-0.65)
Non-EHE						
Midwest	572 (16.61)	481 (18.69)	91 (10.39)	2.13 (1.65-2.76)	—	1.13 (0.81-1.57)
South	358 (10.39)	293 (11.39)	65 (7.47)	1.82 (1.35-2.45)	—	1.19 (0.80-1.78)
Northeast	291 (8.45)	235 (9.13)	56 (6.42)	1.69 (1.23-2.33)	—	0.96 (0.63-1.45)
West	226 (6.56)	173 (6.72)	53 (6.08)	1.32 (0.94-1.84)	—	0.66 (0.43-1.03)
HIV status, n (%)					<.001	—
HIV negative	2561 (74.36)	2228 (86.59)	333 (38.18)	Reference		
PrEP ^f or PEP ^g	769 (22.33)	313 (12.16)	456 (52.36)	0.10 (0.09-0.12)		
HIV positive	114 (3.31)	32 (1.24)	82 (9.34)	0.06 (0.04-0.09)		
Number of risk criteria^h, n (%)					<.001	
2 or more	1241 (36.03)	756 (29.38)	485 (55.73)	Reference		Reference
Less than 2	2203 (63.97)	1817 (70.62)	386 (44.34)	3.02 (2.58-3.54)		1.35 (1.08-1.69)
PrEP indicatedⁱ, n (%)					.52	—

Characteristics	Total	HIV self-test	Other status verification	Univariate OR ^a (95% CI)	<i>P</i> value ^b	Multivariate OR (95% CI)
Yes	2972 (86.29)	2226 (86.51)	746 (85.56)	Reference		
No	472 (13.7)	347 (13.53)	125 (14.36)	0.93 (0.75-1.16)		
Substance use^j, n (%)					.003	
Yes	533 (15.48)	365 (18.58)	168 (22.78)	Reference		Reference
No	2133 (61.93)	1597 (81.36)	536 (76.13)	1.37 (1.11-1.69)		0.77 (0.60-1.00)

^aOR: odds ratio.

^bResults of logistic regression (outcome variable=HIV self-test vs. other HIV verification).

^cNot available.

^dDenotes missing data.

^eEHE: ending the HIV Epidemic priority jurisdictions [15].

^fPrEP: pre-exposure prophylaxis.

^gPEP: postexposure prophylaxis.

^hRisk criteria: (a) inconsistent condom use, (b) HIV-positive partner, and/or (c) bacterial sexually transmitted infection in the past 6 months.

ⁱPre-exposure prophylaxis indication is based on Centers for Disease Control and Prevention's draft of clinical practice guidelines [16]: HIV negative, recent anal or vaginal sex, and one of the following: inconsistent condom usage, HIV-positive partner, or bacterial sexually transmitted infection in past 6 months.

^jSubstance use: cocaine, meth, or heroin use in past 6 months.

Adjunct Surveys

Response rates to adjunct survey measures were mixed, with some successful surveys leading to independent publications [12,13]. Over half of the participants in the cohort responded to the COVID-19 survey (2182/3444, 67.26%) [13]. The Truvada lawsuit survey was offered to HIV-negative participants in the cohort, and again, over half (1485/2555, 58.12%) of those eligible participated [12]. The survey of eligible youth who did not enroll had a much lower response rate, with 33 (33/246, 13.4%) of participants completing or partially completing the survey of the 246 participants who were invited to participate. The mean age of participants was 18.6 years (SD 1.5). The majority of participants were cisgender men (21/33, 64%), White (14/33, 42%), gay (13/33, 39%), had completed high school (27/33, 82%), and knew their HIV status (30/33, 91%). When asked about their number one barrier to participate, the most common response was the HIV testing requirement (13/28, 46%), followed by the length of time to receive the test kit and payment (6/28, 21%), the length of the survey (4/28, 14%), and the fact that the study used a website rather than an app (3/28, 11%). HIV testing was particularly a barrier for younger participants, with a majority (9/11, 82%) of younger participants reporting HIV testing as their number one barrier, compared with about one quarter (5/21, 24%) of older participants. When asked about the best platform to advertise the study on, 58% (18/31) of participants reported Instagram to be the best platform, followed by Twitter (5/31, 16%) and Tiktok (3/31, 10%). Most participants agreed (29/30, 97%) that US \$20 was an appropriate compensation for a short survey. Many participants expressed some privacy concerns, including the privacy of the HIV testing results (12/29, 41%), privacy in testing, emailing, and receiving mail from staff (11/30, 37%), not being out at home as LGBTQ+ (9/29, 31%), and the privacy of the web-based survey (3/29, 10%).

Discussion

Principal Findings

This evaluation of participants enrolled in the Keeping it LITE study indicates that electronic methods are feasible for recruitment of a large, diverse sample of youth and young adults at risk for HIV. We recruited participants representing a broad range of racial, ethnic, gender, and economic groups, although there was limited success in recruiting minor participants. The original goal of recruiting a sample that contained 20% of participants under the age of 18 was not achieved, as the final sample included only 5.95% (205/3444) participants aged <18 years. The racial identity of participants (Table 2) was similar to the 2019 US census estimates [17], although the sample included a slightly smaller proportion of White individuals. Recruitment of transgender and genderqueer participants was an area of success for this study, especially with the large number of trans men recruited, a notably understudied population in HIV research [18,19]. Finally, the study was successful in recruiting participants who were recently diagnosed with HIV, as well as those who were using PrEP and PEP. The rates of PrEP use among HIV-negative participants in this study (769/3330, 23.09%) were comparable with recent estimates of PrEP use compiled by the National HIV Behavioral Surveillance Study [20].

The Keeping it LITE study was 1 of 4 studies funded by the National Institutes of Health to investigate the implementation of large-scale digital HIV interventions for sexual minority men and transgender persons [21-23]. All studies have found success in recruiting large, web-based samples of individuals at risk for HIV, including transgender women [21] and sexual minority men [22]. However, the Keeping it LITE study was able to recruit a larger sample of minor youth compared with other studies recruiting youth [22,23] and was the only study to recruit youth aged 13-15 years. In addition, the only other study that

included transgender men [23] recruited a much smaller sample ($n=53$) than this study ($n=277$). Thus, although this study did not meet the recruitment goal for minor participants, recruitment of minors and transgender men remains a unique contribution of these methods to the literature.

YAB Successes

The formation and use of YAB was particularly successful in this study. The success in formation was in part because of having an established YAB at the sponsoring institution. The researchers were able to build on the established structure with additional youth advisors from around the country. In addition, the researchers used the community guidelines and resources established by the National Institute of Allergy and Infectious Diseases–funded HIV/AIDS Network Collaboration and trained staff to be *adult accomplices*, learning to work alongside youth as equal partners. On the basis of this experience, researchers recommend using and building off existing community engagement structures and providing research staff communications training to address health issues among diverse youth.

Differences Between Recruitment Methods

Most participants were recruited through social media advertisements (1684/2649, 63.57%), although other methods were notably effective for recruiting specific populations. For instance, although word-of-mouth recruitment only accounted for about 18.83% (499/2649) of overall recruitment, this method was vital in the recruitment of Black and trans women participants, leading to the recruitment of nearly half of the trans women and one-third of Black participants. In addition, LGBTQ+ apps only recruited about 12.87% (341/2649) of the overall sample but allowed for the recruitment of individuals with a higher risk profile than other methods. Finally, having a variety of recruitment methods was vital to bolster the diversity of the sample.

Differences in Enrollment Across Demographics

Demographic disparities in enrollment highlighted groups that may face greater barriers to study participation. There were lower enrollment rates for participants who were genderqueer, minors, persons of color (excluding Asian participants), and not using PrEP or PEP. Of concern, many of these demographic groups, including Black and Hispanic or Latinx individuals and those not using PrEP or PEP, are the groups with the greatest burden of HIV risk. For instance, Black and Latino sexual minority men accounted for 38% and 33% of new HIV diagnoses among sexual minority men in 2019, respectively [24], but of the sexual minority men in this study's sample, only 13% were Black and 21% were Latinx. Although this study was not able to recruit a sample that is fully representative of those with new HIV infections, the sample did match US racial or ethnic demographics, and included a large number of gender-diverse participants. The sample also contained a representative number of PrEP and PEP users based on recent data on PrEP uptake among sexual minority men [25]. This speaks to the efforts made to be inclusive and widespread in the advertisement for the study.

Feasibility of At-home Testing

At-home self-testing, a relatively novel outcome assessment method for HIV research trials, proved to be an acceptable, and in fact preferred, method for this sample. The only demographic group in which the majority did not test with an at-home testing kit were PrEP and PEP users. This is likely because they were getting HIV tested as part of their care. In particular, young participants were much more likely to use at-home testing, with nearly all minor participants opting to use this method. Trans men were also more likely to use at-home testing kits, likely because far fewer trans men in this sample used PrEP or PEP.

Feasibility of Adjunct Surveys

Of the adjunct surveys administered during the course of this study, surveys assessing the impact of the Truvada lawsuit and COVID-19 pandemic proved to be easy to administer and provided useful data quickly on large numbers of participants. The adjunct survey administered to youth who did not enroll was less successful because of limited participation. This was an understandable outcome, given that participants who failed to engage in the study were likely to also be less engaged with a follow-up survey. However, it should be noted that this study design allowed for rapid administration of timely surveys, and participant response, when surveying enrolled participants, was efficient and effective.

Adaptations to Barriers in Recruitment of Minor Participants

Minor participants proved particularly challenging to recruit for this study because (1) they tended to have fewer HIV risk behaviors and were therefore less likely to qualify, (2) they were less likely to enroll if eligible, and (3) there were difficulties in advertising to minors on some social media platforms (ie, Snapchat). These barriers were addressed by (1) lowering the risk profile required to participate for minors, (2) shifting social media advertising to platforms that allowed advertising to minors and were commonly used by youth (ie, Instagram), and (3) adjusting advertisement content to avoid being flagged as *inappropriate*. Those who responded to the survey of nonparticipant youth indicated that the most common barrier to participating was the HIV testing requirement and the logistics around it (ie, time to receive kit and privacy concerns). The challenges of recruiting minor participants for this study reflected the researchers' previous experience that younger participants need a larger number of touchpoints in longitudinal studies: more frequent personal interaction, more frequent reminders, and more practical help with study procedures than older participants [26].

Conclusions

The results of enrollment for the Keeping it LITE study indicate that limited interaction recruitment is feasible and well accepted among youth and young adults at risk for HIV. The study was able to quickly enroll a large cohort that broadly reflects the demographics of HIV infections in the US, though enriched for trans- and gender-nonconforming individuals. Several significant barriers to the limited interaction enrollment of adolescent participants were identified. Although some of these barriers were overcome with creative recruitment methods and input

from the YAB, remote enrollment, and particularly at-home HIV testing, are not likely to be successful in enrolling minors at risk for HIV.

Although this cohort study was designed to specifically address HIV incidence and prevention, the framework developed for this study could also be used to investigate social and behavioral epidemiological research questions for sexual and gender minority youth that go beyond HIV. It could also be adapted to use mixed methods, such as incorporating survey questionnaires with open-ended questions, and recruiting a large enough cohort to power inter- and intracategorical intersectional analyses. Regardless of future adaptations, the current implementation demonstrated that consistency is important for retention, such

as regular communication with participants, adjusting procedures in response to participant feedback, and updating study incentives to align with participant preferences.

Going forward, this unique cohort will provide invaluable data to inform prevention strategies and to articulate the best methodologies for limited interaction research. We intend to meticulously characterize factors that put youth and young adults at risk for HIV but also describe the personal and environmental characteristics associated with healthy behaviors and successful choices as participants mature. Our aims include thorough characterization of PrEP attitudes, uptake and adherence, and the predictors of successful navigation through the HIV continuum of care for youth.

Acknowledgments

The authors would like to acknowledge the funders, the National Institutes of Allergy and Infectious Disease (UG3 AI 133676 & UH3 AI 133676), National Institute of Mental Health, the Eunice Kennedy Shriver National Institutes of Child Health and Human Development, their institutions at Cook County Health and the Hektoen Institute of Medicine, and the study participants of the Keeping it LITE study. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

NG and PS drafted the article, including conducting the primary quantitative analyses; PS, AM, AF, and SH conducted data collection; and AF and SH approved the study design concept and findings. All authors reviewed and revised the manuscript.

Conflicts of Interest

None declared.

References

1. Sexually transmitted disease surveillance 2018. Centers for Disease Control and Prevention. 2019. URL: <https://stacks.cdc.gov/view/cdc/79370> [accessed 2021-08-29]
2. Wilson EC, Garofalo R, Harris RD, Herrick A, Martinez M, Martinez J, Transgender Advisory Committee the Adolescent Medicine Trials Network for HIV/AIDS Interventions. Transgender female youth and sex work: HIV risk and a comparison of life factors related to engagement in sex work. *AIDS Behav* 2009 Oct;13(5):902-913 [FREE Full text] [doi: [10.1007/s10461-008-9508-8](https://doi.org/10.1007/s10461-008-9508-8)] [Medline: [19199022](https://pubmed.ncbi.nlm.nih.gov/19199022/)]
3. Garofalo R, Deleon J, Osmer E, Doll M, Harper GW. Overlooked, misunderstood and at-risk: exploring the lives and HIV risk of ethnic minority male-to-female transgender youth. *J Adolesc Health* 2006 Mar;38(3):230-236. [doi: [10.1016/j.jadohealth.2005.03.023](https://doi.org/10.1016/j.jadohealth.2005.03.023)] [Medline: [16488820](https://pubmed.ncbi.nlm.nih.gov/16488820/)]
4. Friedman MR, Wei C, Klem ML, Silvestre AJ, Markovic N, Stall R. HIV infection and sexual risk among men who have sex with men and women (MSMW): a systematic review and meta-analysis. *PLoS One* 2014 Jan 30;9(1):e87139 [FREE Full text] [doi: [10.1371/journal.pone.0087139](https://doi.org/10.1371/journal.pone.0087139)] [Medline: [24498030](https://pubmed.ncbi.nlm.nih.gov/24498030/)]
5. Steinberg L. A social neuroscience perspective on adolescent risk-taking. In: *Biosocial Theories of Crime*. Oxfordshire England: Routledge; 2010.
6. Casey B, Galvan A, Hare TA. Changes in cerebral functional organization during cognitive development. *Curr Opin Neurobiol* 2005 Apr;15(2):239-244. [doi: [10.1016/j.conb.2005.03.012](https://doi.org/10.1016/j.conb.2005.03.012)] [Medline: [15831409](https://pubmed.ncbi.nlm.nih.gov/15831409/)]
7. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. *Am Psychol* 2000 May;55(5):469-480. [Medline: [10842426](https://pubmed.ncbi.nlm.nih.gov/10842426/)]
8. Arnett J. Reckless behavior in adolescence: a developmental perspective. *Dev Rev* 1992 Dec;12(4):339-373. [doi: [10.1016/0273-2297\(92\)90013-r](https://doi.org/10.1016/0273-2297(92)90013-r)]
9. Harper G, Schneider M. Oppression and discrimination among lesbian, gay, bisexual, and Transgendered people and communities: a challenge for community psychology. *Am J Community Psychol* 2003 Jun;31(3-4):243-252. [doi: [10.1023/a:1023906620085](https://doi.org/10.1023/a:1023906620085)] [Medline: [12866682](https://pubmed.ncbi.nlm.nih.gov/12866682/)]
10. Facebook has blocked ad targeting by sexual orientation. Buzz Feed. 2018. URL: <https://www.buzzfeednews.com/article/alexkantrowitz/facebook-has-blocked-ad-targeting-by-sexual-orientation> [accessed 2021-05-26]
11. State laws that enable a minor to provide informed consent to receive HIV and STD services. Centers for Disease Control and Prevention. 2021. URL: <https://www.cdc.gov/hiv/policies/law/states/minors.html> [accessed 2021-05-26]

12. Serrano P, Daubert E, Munoz A, Hosek S, French A. Effect of Truvada lawsuit advertising on preexposure prophylaxis attitudes and decisions among sexual and gender minority youth and young adults at risk for HIV. *AIDS* 2021 Jan 01;35(1):131-139. [doi: [10.1097/QAD.0000000000002728](https://doi.org/10.1097/QAD.0000000000002728)] [Medline: [33031102](https://pubmed.ncbi.nlm.nih.gov/33031102/)]
13. Serrano P, Diaz J, Munoz A, Hosek S, French A. Social, economic, mental health and medical care impacts of COVID-19 in a US cohort of sex and gender diverse adolescents and young adults at-risk for HIV. *J Int AIDS Soc* 2021 Jan;24(S1):33.
14. JASP 0.14.1 – Minor Update. JASP. 2018. URL: <https://jasp-stats.org/2020/12/17/jasp-0-14-1-minor-updates/> [accessed 2021-08-29]
15. Ending the HIV epidemic. Centers for Disease Control and Prevention. 2021. URL: <https://www.cdc.gov/endinghiv/jurisdictions.html> [accessed 2021-05-26]
16. Preexposure prophylaxis for the prevention of HIV infection in the united states– 2021 update. US Public Health service. 2021. URL: <https://www.hiv.gov/blog/cdc-seeks-input-draft-update-prep-guidelines> [accessed 2021-08-29]
17. QuickFacts: United States. United States Census Bureau. URL: <https://www.census.gov/quickfacts/fact/table/US/PST045219> [accessed 2021-05-26]
18. Clark H, Babu AS, Wiewel EW, Opoku J, Crepaz N. Diagnosed HIV infection in transgender adults and adolescents: results from the national HIV surveillance system, 2009-2014. *AIDS Behav* 2017 Sep;21(9):2774-2783 [FREE Full text] [doi: [10.1007/s10461-016-1656-7](https://doi.org/10.1007/s10461-016-1656-7)] [Medline: [28035497](https://pubmed.ncbi.nlm.nih.gov/28035497/)]
19. HIV and transgender People. Centers for Disease Control and Prevention. 2021. URL: <https://www.cdc.gov/hiv/group/gender/transgender/index.html> [accessed 2021-05-26]
20. Finlayson T, Cha S, Xia M, Trujillo L, Denson D, Prejean J, National HIV Behavioral Surveillance Study Group. Changes in HIV preexposure prophylaxis awareness and use among men who have sex with men - 20 urban areas, 2014 and 2017. *MMWR Morb Mortal Wkly Rep* 2019 Jul 12;68(27):597-603 [FREE Full text] [doi: [10.15585/mmwr.mm6827a1](https://doi.org/10.15585/mmwr.mm6827a1)] [Medline: [31298662](https://pubmed.ncbi.nlm.nih.gov/31298662/)]
21. Nash D, Stief M, MacCrate C, Mirzayi C, Patel VV, Hoover D, et al. A web-based study of HIV prevention in the era of pre-exposure prophylaxis among vulnerable HIV-negative gay and bisexual men, transmen, and transwomen who have sex with men: protocol for an observational cohort study. *JMIR Res Protoc* 2019 Sep 17;8(9):e13715 [FREE Full text] [doi: [10.2196/13715](https://doi.org/10.2196/13715)] [Medline: [31538945](https://pubmed.ncbi.nlm.nih.gov/31538945/)]
22. Rendina HJ, Talan AJ, Tavella NF, Matos JL, Jimenez RH, Jones SS, et al. Leveraging technology to blend large-scale epidemiologic surveillance with social and behavioral science methods: successes, challenges, and lessons learned implementing the UNITE longitudinal cohort study of HIV risk factors among sexual minority men in the United States. *Am J Epidemiol* 2021 Apr 06;190(4):681-695 [FREE Full text] [doi: [10.1093/aje/kwaa226](https://doi.org/10.1093/aje/kwaa226)] [Medline: [33057684](https://pubmed.ncbi.nlm.nih.gov/33057684/)]
23. Wirtz AL, Cooney EE, Stevenson M, Radix A, Poteat T, Wawrzyniak AJ, American Cohort To Study HIV Acquisition Among Transgender Women (LITE) Study Group. Digital epidemiologic research on multilevel risks for HIV acquisition and other health outcomes among transgender women in eastern and southern United States: protocol for an online cohort. *JMIR Res Protoc* 2021 Apr 26;10(4):e29152 [FREE Full text] [doi: [10.2196/29152](https://doi.org/10.2196/29152)] [Medline: [33900202](https://pubmed.ncbi.nlm.nih.gov/33900202/)]
24. U.S. Statistics. HIV.org. 2021. URL: <https://www.hiv.gov/hiv-basics/overview/data-and-trends/statistics> [accessed 2021-07-03]
25. Finlayson T, Cha S, Xia M, Trujillo L, Denson D, Prejean J, National HIV Behavioral Surveillance Study Group. Changes in HIV Preexposure Prophylaxis Awareness and Use Among Men Who Have Sex with Men - 20 Urban Areas, 2014 and 2017. *MMWR Morb Mortal Wkly Rep* 2019 Jul 12;68(27):597-603 [FREE Full text] [doi: [10.15585/mmwr.mm6827a1](https://doi.org/10.15585/mmwr.mm6827a1)] [Medline: [31298662](https://pubmed.ncbi.nlm.nih.gov/31298662/)]
26. Hosek SG, Landovitz RJ, Kapogiannis B, Siberry GK, Rudy B, Rutledge B, et al. Safety and feasibility of antiretroviral preexposure prophylaxis for adolescent men who have sex with men aged 15 to 17 years in the United States. *JAMA Pediatr* 2017 Nov 01;171(11):1063-1071 [FREE Full text] [doi: [10.1001/jamapediatrics.2017.2007](https://doi.org/10.1001/jamapediatrics.2017.2007)] [Medline: [28873128](https://pubmed.ncbi.nlm.nih.gov/28873128/)]

Abbreviations

LGBTQ: lesbian, gay, bisexual, transgender, and queer

OR: odds ratio

PEP: postexposure prophylaxis

PrEP: pre-exposure prophylaxis

STI: sexually transmitted infection

YAB: youth advisory board

Edited by G Eysenbach; submitted 27.05.21; peer-reviewed by B Wilbourn, S Lebu; comments to author 28.06.21; revised version received 09.07.21; accepted 01.08.21; published 26.11.21.

Please cite as:

Gleason N, Serrano PA, Muñoz A, French A, Hosek S

Limited Interaction Targeted Epidemiology of HIV in Sexual and Gender Minority American Adolescents and Adults: Feasibility of the Keeping it LITE Study

JMIR Form Res 2021;5(11):e30761

URL: <https://formative.jmir.org/2021/11/e30761>

doi: [10.2196/30761](https://doi.org/10.2196/30761)

PMID: [34346403](https://pubmed.ncbi.nlm.nih.gov/34346403/)

©Neil Gleason, Pedro A Serrano, Alejandro Muñoz, Audrey French, Sybil Hosek. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.11.2021. 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.

Review

Usability-In-Place—Remote Usability Testing Methods for Homebound Older Adults: Rapid Literature Review

Jordan R Hill^{1,2}, PhD; Janetta C Brown^{2,3,4}, PhD; Noll L Campbell^{1,2,4}, PharmD; Richard J Holden^{2,4,5}, PhD

¹Department of Pharmacy Practice, College of Pharmacy, Purdue University, Indianapolis, IN, United States

²Indiana University Center for Aging Research, Regenstrief Institute, Indianapolis, IN, United States

³School of Medicine, Indiana University, Indianapolis, IN, United States

⁴Center for Health Innovation and Implementation Science, Indiana University, Indianapolis, IN, United States

⁵Department of Health and Wellness Design, School of Public Health-Bloomington, Indiana University, Bloomington, IN, United States

Corresponding Author:

Jordan R Hill, PhD

Department of Pharmacy Practice

College of Pharmacy

Purdue University

640 Eskenazi Ave

Indianapolis, IN, 46202

United States

Phone: 1 7655438559

Email: hill265@purdue.edu

Abstract

Background: Technology can benefit older adults in many ways, including by facilitating remote access to services, communication, and socialization for convenience or out of necessity when individuals are homebound. As people, especially older adults, self-quarantined and sheltered in place during the COVID-19 pandemic, the importance of usability-in-place became clear. To understand the remote use of technology in an ecologically valid manner, researchers and others must be able to test usability remotely.

Objective: Our objective was to review practical approaches for and findings about remote usability testing, particularly remote usability testing with older adults.

Methods: We performed a rapid review of the literature and reported on available methods, their advantages and disadvantages, and practical recommendations. This review also reported recommendations for usability testing with older adults from the literature.

Results: Critically, we identified a gap in the literature—a lack of remote usability testing methods, tools, and strategies for older adults, despite this population's increased remote technology use and needs (eg, due to disability or technology experience). We summarized existing remote usability methods that were found in the literature as well as guidelines that are available for conducting in-person usability testing with older adults.

Conclusions: We call on the human factors research and practice community to address this gap to better support older adults and other homebound or mobility-restricted individuals.

(*JMIR Form Res* 2021;5(11):e26181) doi:[10.2196/26181](https://doi.org/10.2196/26181)

KEYWORDS

mobile usability testing; usability inspection; methods; aging; literature synthesis; usability study; mobile usability; elderly; older adults; remote usability; mobility restriction

Introduction

The Need for Remote Operations

Technology can support access to services, communication, and socialization for older adults and others whose mobility is

restricted due to health-related risks such as susceptibility to disease (eg, COVID-19), disability, and a lack of resources (eg, transportation). However, the delivery, support, and evaluation of technologies that are used by homebound or mobility-restricted individuals require remote operations,

including remote usability testing. Herein, we review the unique needs and technology opportunities of homebound older adults and the literature on remote usability testing methods. Based on our findings, we identified a gap in guidance for remote usability testing with older adults. Therefore, we call on relevant research and practice communities to address this gap.

Supporting Homebound Individuals With Technology

Over 2 million Americans are homebound due to an array of social, functional, and health-related causes, and this number is projected to grow as the size of the older population increases [1]. Situational factors such as inclement weather and, on a larger scale, pandemics or national disasters can also temporarily render individuals homebound. For example, in March 2020, the US Centers for Disease Control and Prevention [2] warned older adults to remain at home due to the disproportionate COVID-19-related health risks that they face. Prior to the pandemic, an estimated 1 in 4 older US adults were already socially isolated, and this rate has likely increased [3].

People who shelter in place or stay home for other reasons may turn to technology to access remote services, including remote banking, grocery shopping, and medical care services. The prevalence of these physically distant interactions is reportedly on the rise [4], especially for certain services. A prominent example is the increased frequency of patients' telemedicine visits with health care professionals—a form of telehealth that has been long available but whose usage has increased dramatically in the United States, as the COVID-19 pandemic resulted in changes to federal reimbursement policies in March 2020 [5].

Testing Technology With Homebound Technology Users

When technology users are homebound, researchers and care practitioners who intend to test a technology's usability in an ecologically valid manner must either travel to the user's home or conduct remote testing. Travel is not always an option. Safety,

health, or personal reasons may prevent researchers from entering a home or community. Travel may be too costly or otherwise impractical, or participants may live in an area that is inaccessible to the project team. During the COVID-19 pandemic for example, academic and practice-based project teams have anecdotally reported barriers to in-person visits, including the need to distance themselves from infected and at-risk individuals, members of project teams working from home, and the need to reduce travel expenses due to economic pressures. Even if in-person visits are possible, remote testing can also be more convenient and cost-efficient for all parties involved.

Methods

We performed a rapid review of studies involving remote usability testing methods for all users and those specifically for older adults and summarized their findings. Rapid reviews are an accepted knowledge synthesis approach that has become popular for understanding the most salient points on emerging or timely topics [6]. Rapid reviews typically do not include an exhaustive set of studies, do not involve formal analyses of study quality, and report findings from prior studies via narrative synthesis [7]. The primary goal of this review was to identify methods for performing remote usability assessments with older adults (if any existed). Secondarily, we wished to summarize the literature on existing remote usability methodologies for any population and existing guidelines on performing in-person usability testing with older adults. Sources for the second goal were largely retrieved while searching for sources to support the primary goal and via a secondary search within Google Scholar.

Our rapid review began with a keyword search on the Google Scholar and Science Direct scholarly databases. This was followed by a supplementary keyword search in top human factors journals and proceedings. Both searches are summarized in [Table 1](#).

Table 1. Keywords that were searched for the rapid review.

Search type and sources	Keywords
Primary search	
Google Scholar (database)	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Science Direct (database)	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Secondary search	
Ergonomics	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Human Factors	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Applied Ergonomics	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Human Factors and Ergonomics Society Conference Proceedings	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
International Journal of Human-Computer Interaction	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
International Journal of Human-Computer Studies	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Gerontechnology	<i>elderly remote usability, senior remote usability, and older adult remote usability</i>
Google Scholar (database)	<i>usability older adults, elderly usability, senior usability, and remote usability</i>

We began with Google Scholar to take advantage of its relevance-based sorting feature and broader inclusion of diverse disciplines, academic and practice-based publications, and grey literature [8]. However, we conducted further searches because of the known limitations of Google Scholar, such as its lack of transparency and lack of specialization [9].

In the interest of establishing a starting point for understanding remote usability testing with older adults, we had broad inclusion criteria and did not restrict studies based on their date of publication or an analysis of their quality or peer-review status. We also defined *remote usability* broadly as usability assessments of participants (users) who were in separate locations from the researchers or practitioners. Duplicate studies, as well as studies in which usability was assessed by an expert (eg, heuristic analysis on a website) on behalf of older adults instead of through direct participant feedback, were excluded.

Two authors (JRH and JCB) performed the search in Google Scholar while one author (JCB) performed the search in Science Direct and the human factors sources. Both authors took notes in a shared cloud-based document. We chose a stopping rule based on the assumption that a narrative synthesis of literature is a form of qualitative content analysis [10]. Therefore, we concluded our search when we reached theoretical saturation [11]—a qualitative analysis stopping rule that means that the search continues until results begin to repeat and negligible new categories of information are produced through additional searching.

Results and Discussion

Summary of the Search Results

Of all of the sources found, 33 were screened in-depth (18 on remote usability methods and 15 on usability testing with older adults), and 21 were included in this review (16 on various remote usability methods and 5 on usability testing with older adults).

Importantly, sources that provided guidance or information on remote usability testing with older adults (the primary goal of this review) were not found. Therefore, we organized the results according to our secondary goals—summarizing existing methods for remote usability testing and outlining existing guidelines for in-person usability testing with older adults. In this *Results and Discussion* section, we combined the results with our interpretations and discussion to adhere to conventions for narrative reviews. We also present our overall conclusions in the *Conclusions* section.

Usability-In-Place: The State of the Practice of Remote Usability Testing

Studies on remote usability testing date back to the 1990s [12,13]. Since then, most traditional in-person usability evaluation methods have been attempted remotely. Remote moderated testing has been supported by advances in internet-based software, such as WebEx and NetMeeting, which permit simultaneous video and audio transmissions, screen sharing, and remote control [14,15]. Studies have also used novel methods, such as using virtual reality to simulate laboratory usability testing environments [16] and remotely

capturing eye-tracking data [17]. Technologies for unmoderated testing have also evolved, as described elsewhere [18].

Asynchronous methods have long been used to overcome the barriers of time and space. Such methods include conducting self-administered survey questionnaires, using user diaries and incident reports, and obtaining voluntary feedback [19]. Studies have also used activity logging to passively collect use data for analyzing usability [20].

The major remote usability testing methods are described in Table 2 along with key findings from the literature. An important replicated finding was that the results from remote and in-person usability testing were generally similar, although significant differences may have appeared under extenuating circumstances, such as poor product usability or the cognitive difficulty of the usability testing tasks [21].

Table 2. Remote usability testing methods and key findings.

Remote usability testing method	Description	Key findings
Synchronous remote testing [14,15,20-23]	In-person testing is simulated by using video and audio transmissions and remote desktop access.	<ul style="list-style-type: none"> Nearly identical to conventional in-person testing (with comparable results) [14,21-23] Indirect cues and context can be missed [20] Participants can prefer remote testing to in-person testing [22] Management challenges (eg, network issues, remote troubleshooting, and setup) [15,20,22] Users take longer to complete tasks than during in-person testing [15] Users make more errors than during in-person testing [15]
Web-based questionnaires or surveys [14,20,21]	Users fill out web-based questionnaires as they complete tasks or after the completion of tasks.	<ul style="list-style-type: none"> More time-consuming for users^a [14] Less time-consuming for users than lab-based usability testing when usability is poor^a [21] Overall usability rated lower when compared to lab-based usability testing [21] Identifies fewer specific usability problems [14] Enables the collection of data from many participants [20] Validity problems with the self-report approach [20]
Postuse interview [24]	Users are interviewed over the phone about the usability of a design (qualitative and quantitative data are collected) after they have completed tasks.	<ul style="list-style-type: none"> Beneficial for those with disabilities Quantitative data collected are comparable to in-person testing data Qualitative data are less rich compared to in-person testing data In-person testing is better for formative testing; remote testing is better for summative evaluation
User-reported critical incidents or diaries [12,13,19,20]	Users fill out a diary and take notes during a period of use or fill out an incident form when they identify a critical problem with an interface.	<ul style="list-style-type: none"> Able to capture most high- and moderate-severity incidents^a [12,13] Users report fewer low-severity incidents than experts [12,13] Validity problems with self-reports [20] Issues may be underreported compared to those reported via traditional methods^a [19]
User-provided feedback [25]	While completing timed tasks, users provide comments or feedback in a separate browser window. Once a task is complete, the user rates the difficulty of the completed task.	<ul style="list-style-type: none"> The percentage of participants who completed remote testing tasks was the same as the percentage of participants who completed in-person testing tasks No difference in the time taken to complete tasks Able to capture rich qualitative information through typed comments Less observation data captured compared to those captured during in-person testing Captured fewer usability issues in some cases compared to those captured during in-person testing
Log analysis [20]	The actions taken by the user (eg, clicks) are captured for future analysis.	<ul style="list-style-type: none"> Less intrusive to user Can collect data from many users Unable to capture user intentions or additional context

^aConflicting evidence has been found to support both the statement and its opposite in the literature.

The following general benefits of remote usability testing methods were identified:

- Does not require a facility, thereby reducing the time requirements of participants and evaluators and lowering costs [20]

- Can recruit participants from a broader geographic vicinity, thereby allowing evaluators to collect results from a larger and more diverse group of people (including those living in other countries or rural areas or those who are otherwise isolated) [14,23]

- Allows participants to test technologies in a more realistic environment. For example, Petrie et al [24] had people with disabilities perform remote usability testing from the comfort of their own homes. The benefits thereof include the use of a home-based environment that is almost impossible to perfectly replicate in a lab.

Several drawbacks were also described, as follows:

- General agreement that remotely collecting data results in a loss of some of the contextual information and nonverbal cues from participants that are collected during in-person evaluations [15,20,22,24,25]
- Remote usability methods (especially asynchronous methods) appear to result in the identification of fewer usability problems, cause users to make more errors during testing, and are more time-consuming for users [14,15]. However, test participants identified about as many usability issues as those identified by the evaluators, but the participants' categorization of the identified usability problems were deemed not useful. Contrarily, this was not found by Tullis et al [25] when they compared lab-based usability testing against remote usability testing.
- Dray and Siegel [20] also listed validity problems with self-report methodologies, the inability of log files to distinguish the cause of navigation errors, and management challenges related to troubleshooting network issues and ensuring system compatibility as other drawbacks of remote usability testing.
- Many of the factors that may affect the validity, reliability, or efficiency of remote usability testing have not been scientifically studied [26]. These include factors such as the characteristics of users (eg, age and literacy), the effect of slow or unstable internet, the type of devices being used, and testing tactics (eg, verbal, printed, or on-screen instructions).

No matter the method, remote usability testing also involves challenges to implementing the methods in natural contexts, namely in home and community settings [27,28]. These challenges include recruiting a representative sample, especially among populations that may be less comfortable with using certain technology, have lower literacy, or are mistrustful of

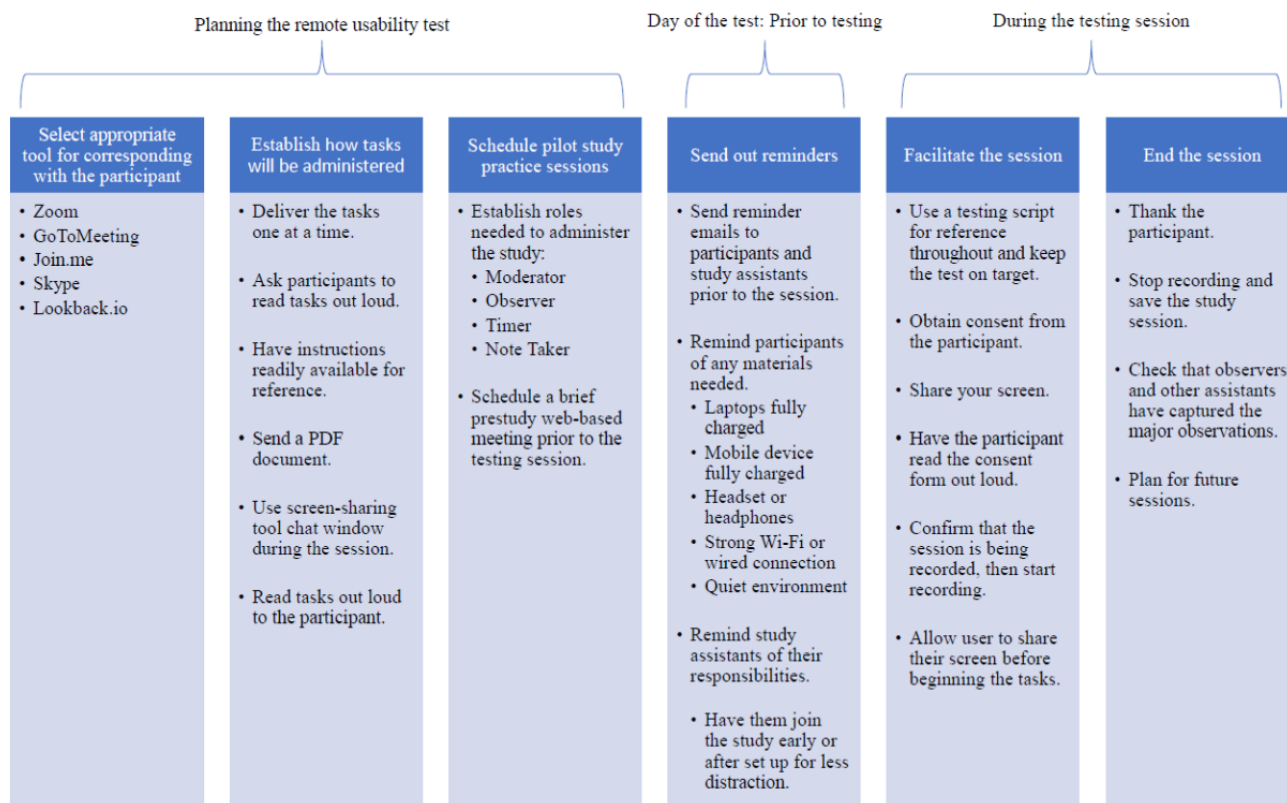
research [26]. McLaughlin and colleagues [26] proposed strategies such as providing access to phone support prior to the start of any web-based testing.

Remote Usability Testing With Older Adults

Prior work on remote usability testing has been performed with convenience samples of college students [13,14] or healthier and younger adults recruited from workplaces [22,23,25]. We found no published instance of fully remote usability testing with older adults. Diamantidis et al [29] conducted a test of a mobile health system with older individuals with chronic kidney disease. Participants received an in-person tutorial of the system; they used the system at home, received physical materials by mail, and completed a paper diary. Afterward, they returned to complete an in-person satisfaction survey. Petrie et al [24] reported 2 case studies of remote usability testing—one with blind younger adults (n=8) and another with a more heterogeneous group of individuals with disabilities (n=51). They demonstrated the feasibility of remote testing and showed comparable results between in-person and remote testing, although in-person participants in the second study reported more usability problems with the tested website.

Others have described ways to improve in-person usability testing with older adults that may be transferable to remote methods. For example, touch screen devices and hardware that is selected for simplicity may produce better usability testing results with older adults [30-32] and can therefore reduce barriers to remote usability testing. Additionally, the use of large closed captions during a remote testing session has been recommended for older users with visual or hearing impairments. Holden [33] published a Simplified System Usability Scale that was modified for and tested with older adults and those with cognitive disabilities but did not demonstrate its use in remote testing.

Older adults in remote usability tests may also benefit from non-age-specific strategies for optimizing remote usability testing [34]. These recommendations, which are summarized in Figure 1, include mailing a written copy of instructions, conducting web-based training prior to testing sessions, and sending reminders.

Figure 1. General guidelines for conducting moderated remote usability testing (adapted from the Nielsen Norman Group [34]).

Conclusions

Our rapid review and synthesis of the literature revealed that remote usability testing still appears to be an emerging field [26] whose great potential is accentuated during major events, such as the COVID-19 pandemic. The decision to pursue the further development of and research on remote usability testing is straightforward, given the apparent advantages, validity, and feasibility of remote usability testing and the need for the method.

The method however must be adapted to and tested with older adults. The use of technology for remote services among older adults in the United States has been increasing [35,36], as has older adults' proficiency with internet-based technology [37]. A Pew Research Center national survey reported increases in internet use (from 12% to 67%) and the adoption of home broadband (from 0% to 51%) from 2000 to 2016, as well as increases in smartphone (from 11% to 42%) and tablet (from 1% to 32%) ownership from 2011 to 2016 [4]. However, the older adult population is diverse and has different needs

compared to those of other groups when it comes to technology and the usability testing of technology. US adults aged 65 years are more likely than their younger counterparts to experience difficulties with physical or cognitive function, including reduced memory capacity, stiff joints or arthritis, and vision or hearing disability [38,39]. These factors and the discomfort with or reduced motivation to use technology elevate the importance of usability testing [40] but ironically may increase the difficulty of conducting remote usability testing. Additional recommendations and best practices will thus be needed to ensure effective and efficient remote usability testing with older adults.

We call on human factors, human-computer interaction, and digital health communities to further develop, describe, and test remote usability testing approaches that will be suitable across diverse populations, including older adults, those with lower literacy or health literacy, and individuals with cognitive or physical disabilities. Progress toward this goal will not only better support homebound or mobility-restricted individuals but may also improve the efficiency, ecological validity, and effectiveness of usability testing in general.

Acknowledgments

We thank Dr Anne McLaughlin and the members of the Brain Safety Lab for their input. The authors of this paper were supported by grant R01 AG056926 from the National Institute on Aging of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the three anonymous reviewers.

Authors' Contributions

JRH and RJH conceived this paper. All authors wrote and edited this paper and approved the final version.

Conflicts of Interest

None declared.

References

1. Ornstein KA, Leff B, Covinsky KE, Ritchie CS, Federman AD, Roberts L, et al. Epidemiology of the homebound population in the United States. *JAMA Intern Med* 2015 Jul;175(7):1180-1186 [FREE Full text] [doi: [10.1001/jamainternmed.2015.1849](https://doi.org/10.1001/jamainternmed.2015.1849)] [Medline: [26010119](https://pubmed.ncbi.nlm.nih.gov/26010119/)]
2. COVID-19 risks and vaccine information for older adults. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/older-adults.html#> [accessed 2020-08-27]
3. Cudjoe TKM, Kotwal AA. "Social Distancing" amid a crisis in social isolation and loneliness. *J Am Geriatr Soc* 2020 Jun;68(6):E27-E29 [FREE Full text] [doi: [10.1111/jgs.16527](https://doi.org/10.1111/jgs.16527)] [Medline: [32359072](https://pubmed.ncbi.nlm.nih.gov/32359072/)]
4. Anderson M, Perrin A. Technology use among seniors. Pew Research Center. 2017 May 17. URL: <https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/> [accessed 2020-11-24]
5. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. *J Am Med Inform Assoc* 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: [10.1093/jamia/ocaa067](https://doi.org/10.1093/jamia/ocaa067)] [Medline: [32311034](https://pubmed.ncbi.nlm.nih.gov/32311034/)]
6. Tricco AC, Antony J, Zarin W, Striffler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. *BMC Med* 2015 Sep 16;13:224 [FREE Full text] [doi: [10.1186/s12916-015-0465-6](https://doi.org/10.1186/s12916-015-0465-6)] [Medline: [26377409](https://pubmed.ncbi.nlm.nih.gov/26377409/)]
7. Grant M, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009 Jun;26(2):91-108 [FREE Full text] [doi: [10.1111/j.1471-1842.2009.00848.x](https://doi.org/10.1111/j.1471-1842.2009.00848.x)] [Medline: [19490148](https://pubmed.ncbi.nlm.nih.gov/19490148/)]
8. Yasin A, Fatima R, Wen L, Afzal W, Azhar M, Torkar R. On using grey literature and Google Scholar in systematic literature reviews in software engineering. *IEEE Access* 2020;8:36226-36243 [FREE Full text] [doi: [10.1109/access.2020.2971712](https://doi.org/10.1109/access.2020.2971712)]
9. Shultz M. Comparing test searches in PubMed and Google Scholar. *J Med Libr Assoc* 2007 Oct;95(4):442-445 [FREE Full text] [doi: [10.3163/1536-5050.95.4.442](https://doi.org/10.3163/1536-5050.95.4.442)] [Medline: [17971893](https://pubmed.ncbi.nlm.nih.gov/17971893/)]
10. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med* 2005 Jul;61(2):417-430. [doi: [10.1016/j.socscimed.2004.12.001](https://doi.org/10.1016/j.socscimed.2004.12.001)] [Medline: [15893056](https://pubmed.ncbi.nlm.nih.gov/15893056/)]
11. Strauss A, Corbin J. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd ed. Thousand Oaks, California: Sage; 1998.
12. Castillo JC, Hartson HR, Hix D. Remote usability evaluation: Can users report their own critical incidents? 1998 Apr Presented at: CHI98: ACM Conference on Human Factors and Computing Systems; April 18-23, 1998; Los Angeles, California, USA p. 253-254. [doi: [10.1145/286498.286736](https://doi.org/10.1145/286498.286736)]
13. Hartson HR, Castillo JC. Remote evaluation for post-deployment usability improvement. 1998 May Presented at: AVI98: Advanced Visual Interface '98; May 24-27, 1998; L'Aquila, Italy p. 22-29. [doi: [10.1145/948496.948499](https://doi.org/10.1145/948496.948499)]
14. Andreasen M, Nielsen H, Schröder S, Stage J. What happened to remote usability testing? An empirical study of three methods. 2007 Presented at: CHI '07: SIGCHI Conference on Human Factors in Computing Systems; April 28 to May 3, 2007; San Jose, California, USA p. 1405-1414. [doi: [10.1145/1240624.1240838](https://doi.org/10.1145/1240624.1240838)]
15. Thompson KE, Rozanski EP, Haake AR. Here, there, anywhere: Remote usability testing that works. 2004 Oct Presented at: SIGITE04: ACM Special Interest Group for Information Technology Education Conference 2004; October 28-30, 2004; Salt Lake City, Utah, USA p. 132-137. [doi: [10.1145/1029533.1029567](https://doi.org/10.1145/1029533.1029567)]
16. Madathil KC, Greenstein JS. Synchronous remote usability testing: a new approach facilitated by virtual worlds. 2011 Presented at: CHI '11: SIGCHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, British Columbia, Canada. [doi: [10.1145/1978942.1979267](https://doi.org/10.1145/1978942.1979267)]
17. Chynał P, Szymański JM. Remote usability testing using eyetracking. 2011 Presented at: IFIP Conference on Human-Computer Interaction – INTERACT 2011; September 5-9, 2011; Lisbon, Portugal p. 356-361 URL: https://doi.org/10.1007/978-3-642-23774-4_29 [doi: [10.1007/978-3-642-23774-4_29](https://doi.org/10.1007/978-3-642-23774-4_29)]
18. Whintont K. Tools for unmoderated usability testing. Nielsen Norman Group. 2019 Sep 22. URL: <https://www.nngroup.com/articles/unmoderated-user-testing-tools/> [accessed 2020-11-24]
19. Bruun A, Gull P, Hofmeister L, Stage J. Let your users do the testing: a comparison of three remote asynchronous usability testing methods. 2009 Apr Presented at: CHI '09: CHI Conference on Human Factors in Computing Systems; April 4-9, 2009; Boston, Massachusetts, USA p. 1619-1628. [doi: [10.1145/1518701.1518948](https://doi.org/10.1145/1518701.1518948)]
20. Dray S, Siegel D. Remote possibilities?: international usability testing at a distance. *Interactions (NY)* 2004;11(2):10-17. [doi: [10.1145/971258.971264](https://doi.org/10.1145/971258.971264)]

21. Sauer J, Sonderegger A, Heyden K, Biller J, Klotz J, Uebelbacher A. Extra-laboratorial usability tests: An empirical comparison of remote and classical field testing with lab testing. *Appl Ergon* 2019 Jan;74:85-96. [doi: [10.1016/j.apergo.2018.08.011](https://doi.org/10.1016/j.apergo.2018.08.011)] [Medline: [30487113](https://pubmed.ncbi.nlm.nih.gov/30487113/)]
22. Bernheim Brush AJ, Ames M, Davis J. A comparison of synchronous remote and local usability studies for an expert interface. 2004 Presented at: CHI04: CHI 2004 Conference on Human Factors in Computing Systems; April 24-29, 2004; Vienna, Austria. [doi: [10.1145/985921.986018](https://doi.org/10.1145/985921.986018)]
23. Hammontree M, Weiler P, Nayak N. Remote usability testing. *Interactions (NY)* 1994 Jul;1(3):21-25. [doi: [10.1145/182966.182969](https://doi.org/10.1145/182966.182969)]
24. Petrie H, Hamilton F, King N, Pavan P. Remote usability evaluations with disabled people. 2006 Presented at: CHI '06: SIGCHI Conference on Human Factors in Computing Systems; April 22-27, 2006; Montréal, Québec, Canada p. 1133-1141. [doi: [10.1145/1124772.1124942](https://doi.org/10.1145/1124772.1124942)]
25. Tullis T, Fleischman S, McNulty M, Cianchette C, Bergel M. An empirical comparison of lab and remote usability testing of Web sites. 2002 Presented at: Usability Professionals' Association 2002 Conference Proceedings; 2002; Bloomington, Illinois, USA.
26. McLaughlin AC, DeLucia PR, Drews FA, Vaughn-Cooke M, Kumar A, Nesbitt RR, et al. Evaluating medical devices remotely: Current methods and potential innovations. *Hum Factors* 2020 Nov;62(7):1041-1060 [FREE Full text] [doi: [10.1177/0018720820953644](https://doi.org/10.1177/0018720820953644)] [Medline: [32960093](https://pubmed.ncbi.nlm.nih.gov/32960093/)]
27. Holden RJ, Scott AMM, Hoonakker PLT, Hundt AS, Carayon P. Data collection challenges in community settings: insights from two field studies of patients with chronic disease. *Qual Life Res* 2015 May;24(5):1043-1055 [FREE Full text] [doi: [10.1007/s11136-014-0780-y](https://doi.org/10.1007/s11136-014-0780-y)] [Medline: [25154464](https://pubmed.ncbi.nlm.nih.gov/25154464/)]
28. Valdez RS, Holden RJ. Health care human factors/ergonomics fieldwork in home and community settings. *Ergon Des* 2016 Oct;24(4):4-9 [FREE Full text] [doi: [10.1177/1064804615622111](https://doi.org/10.1177/1064804615622111)] [Medline: [28781512](https://pubmed.ncbi.nlm.nih.gov/28781512/)]
29. Diamantidis CJ, Ginsberg JS, Yoffe M, Lucas L, Prakash D, Aggarwal S, et al. Remote usability testing and satisfaction with a mobile health medication inquiry system in CKD. *Clin J Am Soc Nephrol* 2015 Aug 07;10(8):1364-1370 [FREE Full text] [doi: [10.2215/CJN.12591214](https://doi.org/10.2215/CJN.12591214)] [Medline: [26220816](https://pubmed.ncbi.nlm.nih.gov/26220816/)]
30. Murata A, Iwase H. Usability of touch-panel interfaces for older adults. *Hum Factors* 2005;47(4):767-776. [doi: [10.1518/001872005775570952](https://doi.org/10.1518/001872005775570952)] [Medline: [16553065](https://pubmed.ncbi.nlm.nih.gov/16553065/)]
31. Page T. Touchscreen mobile devices and older adults: a usability study. *Int J Hum Factors Ergon* 2014;3(1):65. [doi: [10.1504/IJHFE.2014.062550](https://doi.org/10.1504/IJHFE.2014.062550)]
32. Ziefle M, Bay S. How older adults meet complexity: Aging effects on the usability of different mobile phones. *Behav Inf Technol* 2005 Sep;24(5):375-389. [doi: [10.1080/0144929042000320009](https://doi.org/10.1080/0144929042000320009)]
33. Holden RJ. A Simplified System Usability Scale (SUS) for Cognitively Impaired and Older Adults. In: Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care. 2020 Sep 16 Presented at: 2020 International Symposium on Human Factors and Ergonomics in Health Ca; March 8-11, 2020; Toronto, Ontario p. 180-182. [doi: [10.1177/2327857920091021](https://doi.org/10.1177/2327857920091021)]
34. Moran K, Pernice K. Remote moderated usability tests: How to do them. Nielsen Norman Group. 2020 Apr 26. URL: <https://www.nngroup.com/articles/moderated-remote-usability-test/> [accessed 2020-11-24]
35. Baker S, Warburton J, Waycott J, Batchelor F, Hoang T, Dow B, et al. Combatting social isolation and increasing social participation of older adults through the use of technology: A systematic review of existing evidence. *Australas J Ageing* 2018 Sep;37(3):184-193. [doi: [10.1111/ajag.12572](https://doi.org/10.1111/ajag.12572)] [Medline: [30022583](https://pubmed.ncbi.nlm.nih.gov/30022583/)]
36. Lai HJ. Investigating older adults' decisions to use mobile devices for learning, based on the unified theory of acceptance and use of technology. *Interactive Learning Environments* 2018 Nov 21;28(7):890-901. [doi: [10.1080/10494820.2018.1546748](https://doi.org/10.1080/10494820.2018.1546748)]
37. Morrow-Howell N, Galucia N, Swinford E. Recovering from the COVID-19 pandemic: A focus on older adults. *J Aging Soc Policy* 2020;32(4-5):526-535. [doi: [10.1080/08959420.2020.1759758](https://doi.org/10.1080/08959420.2020.1759758)] [Medline: [32336225](https://pubmed.ncbi.nlm.nih.gov/32336225/)]
38. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. *Comput Human Behav* 2013 Jul;29(4):1715-1724. [doi: [10.1016/j.chb.2013.02.006](https://doi.org/10.1016/j.chb.2013.02.006)]
39. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *Int J Med Inform* 2018 Jun;114:66-75. [doi: [10.1016/j.ijmedinf.2018.03.012](https://doi.org/10.1016/j.ijmedinf.2018.03.012)] [Medline: [29673606](https://pubmed.ncbi.nlm.nih.gov/29673606/)]
40. Brown J, Kim HN. Validating the usability of an Alzheimer's caregiver mobile app prototype. 2020 Presented at: 2020 IISE Annual Conference; May 30 to June 2, 2020; New Orleans, Louisiana, USA.

Edited by A Kushniruk; submitted 01.12.20; peer-reviewed by A Federman, K Yin, L Perrier; comments to author 23.12.20; revised version received 25.01.21; accepted 02.09.21; published 02.11.21.

Please cite as:

Hill JR, Brown JC, Campbell NL, Holden RJ

Usability-In-Place—Remote Usability Testing Methods for Homebound Older Adults: Rapid Literature Review

JMIR Form Res 2021;5(11):e26181

URL: <https://formative.jmir.org/2021/11/e26181>

doi: [10.2196/26181](https://doi.org/10.2196/26181)

PMID: [34726604](https://pubmed.ncbi.nlm.nih.gov/34726604/)

©Jordan R Hill, Janetta C Brown, Noll L Campbell, Richard J Holden. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 02.11.2021. 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

Twitter Use by Academic Nuclear Medicine Programs: Pilot Content Analysis Study

Ananya Panda^{1*}, MD; Akash Sharma^{2*}, MBA, MD; Ayca Dundar¹, MD; Ann Packard¹, MD; Lee Aase³, MBA; Amy Kotsenas¹, MD; Ayse Tuba Kendi¹, MD

¹Department of Radiology, Mayo Clinic, Rochester, MN, United States

²Department of Radiology, Mayo Clinic, Jacksonville, FL, United States

³Social Media Network, Mayo Clinic, Rochester, MN, United States

*these authors contributed equally

Corresponding Author:

Ayse Tuba Kendi, MD
Department of Radiology
Mayo Clinic
200 1st Street SW
Rochester, MN, 55905
United States
Phone: 1 612 408 9737
Email: kendi.ayse@mayo.edu

Abstract

Background: There is scant insight into the presence of nuclear medicine (NM) and nuclear radiology (NR) programs on social media.

Objective: Our purpose was to assess Twitter engagement by academic NM/NR programs in the United States.

Methods: We measured Twitter engagement by the academic NM/NR community, accounting for various NM/NR certification pathways. The Twitter presence of NM/NR programs at both the department and program director level was identified. Tweets by programs were cross-referenced against potential high-yield NM- or NR-related hashtags, and tabulated at a binary level. A brief survey was done to identify obstacles and benefits to Twitter use by academic NM/NR faculty.

Results: For 2019-2020, 88 unique programs in the United States offered NM/NR certification pathways. Of these, 52% (46/88) had Twitter accounts and 24% (21/88) had at least one post related to NM/NR. Only three radiology departments had unique Twitter accounts for the NM/molecular imaging division. Of the other 103 diagnostic radiology residency programs, only 16% (16/103) had a presence on Twitter and 5% (5/103) had tweets about NM/NR. Only 9% (8/88) of NM/NR program directors were on Twitter, and three program directors tweeted about NM/NR. The survey revealed a lack of clarity and resources around using Twitter, although respondents acknowledged the perceived value of Twitter engagement for attracting younger trainees.

Conclusions: Currently, there is minimal Twitter engagement by the academic NM/NR community. The perceived value of Twitter engagement is counterbalanced by identifiable obstacles. Given radiologists' overall positive views of social media's usefulness, scant social media engagement by the NM community may represent a missed opportunity. More Twitter engagement and further research by trainees and colleagues should be encouraged, as well as the streamlined use of unique hashtags.

(*JMIR Form Res* 2021;5(11):e24448) doi:[10.2196/24448](https://doi.org/10.2196/24448)

KEYWORDS

social media; Twitter; radiology; nuclear medicine; nuclear radiology; social network; medical education; networking

Introduction

Social media use among adults in the United States has increased substantially over the last decade, with the percentage of people who use at least one social media account doubling from 36% in 2009 to 72% in 2019 [1]. As with the general population

trends, there has been a burgeoning use of social media by health care professionals, including radiologists. Although early research on the use of social media focused on physician-patient communication, the same enthusiasm is now developing for professional networking, education, and peer engagement. In recent surveys, 60%-85% of radiologist respondents

acknowledged using social media for a mix of professional and personal purposes [2,3]. Specifically, social media is more popular among medical students and radiology trainees; they aim to use it for education, information gathering, informal engagement, and mentorship opportunities with experienced faculty members [4]. Of all the social media platforms offering unique features to engage with a wider audience online [5], Twitter is particularly popular for professional use as it allows real-time, multilateral online conversations, and sharing of a wide range of content [6]. Both program directors as individuals and academic radiology departments as group entities are increasingly receptive to Twitter as an effective medium for academic activity; this includes radiology education, recruitment of future faculty and trainees, peer networking, and opportunities for collaboration and engagement on the basis of similar interests [7-9].

As a specialty, nuclear medicine (NM) and Molecular Imaging is complex as there are several pathways to achieve posttraining certification in NM or nuclear radiology (NR). Physicians and scientists—both those trained in radiology and those not trained in radiology—pursue the various training pathway options in the United States [10]. Currently, training pathways include the following: (1) a traditional NM residency program embedded within the radiology department (~39 programs as of 2019), (2) a 1-year NR fellowship after diagnostic radiology training, and (3) a dual pathway comprising 16 months of NR training integrated into a diagnostic radiology (DR) or intervention radiology residency (~50 DR programs and 1 intervention radiology program as of 2019). Despite attempts to increase the number of physicians entering NM/NR by reforming training requirements, a recent survey found a lack of exposure to NM and NR among medical students and early radiology residents [11], with reinforced calls to improve outreach efforts to medical students and radiology residents [11,12].

Although these two developments as described are separate, we propose that social media may represent an opportunity and serve as a resource that the academic NM/NR community can use for outreach efforts aimed at attracting future trainees. However, to our knowledge, there is very little information on the presence of NM and NR programs on social media, specifically Twitter.

Social media analytics tools such as Symplur allow codifying of data through ontology hashtags on Twitter [13]. These ontology hashtags aid in the topical organization of tweets, thereby channeling overwhelming amounts of data into more relevant and consumable data streams. The radiology tag ontology project was devised in 2015 as an initiative specifically aimed at codifying a list of radiology-related hashtags that people can use to tag social media content so that it may be discovered by others with similar interests [13].

Our purpose was to assess Twitter engagement by academic NM/NR programs in the United States, and to further characterize the value proposition of social media use in this specialty.

Methods

Institutional Review Board approval was deemed not necessary for this study as it is an internet-based study accessing publicly available information. We compiled a composite list of programs that offered NM residency, NR fellowships, and dual DR/NR pathways and radiology residencies for 2019-2020, obtained from the American Board of Medical Specialties, the American Board of Nuclear Medicine (ABNM), the Accreditation Council for Graduate Medical Education, and the Electronic Residency Application System [14-16]. If a program offered more than one training pathway, it was counted only once and was grouped into NM residency, NR fellowship, and DR/NR residency in decreasing hierarchical order. The list of program directors for NM residencies, fellowship directors for NR fellowships, program directors for DR programs, and NR division chief preceptors for the 16-month DR/NR pathway was obtained through the same online sources, as well as the American Board of Radiology (ABR); we also used individual radiology program websites for any missing data [15,16].

In December 2019, a manual search for Twitter handles was done for each of these radiology programs, followed by consensus reconciliation by a radiology fellow and a radiology resident. The search process was iterative, using the “Search,” “Top,” and “People” features of Twitter. The Twitter presence of a training program was considered positive if the program had a Twitter handle (account) for the radiology department or radiology residents. If a program had additional Twitter handles specifically for its NM/NR division, it was noted separately. A manual search using the “Search” and “People” features was also done to identify individual Twitter handles of program directors for NM residencies, fellowship directors for NR fellowships, program directors for DR programs, and/or NR division chairs for those programs only offering DR/NR pathways. These individuals are collectively referred to as program directors throughout. A program director Twitter presence was considered positive if any of the above individuals meeting the definition of a program director had an individual Twitter account. For general radiology residency programs that do not offer NM/NR training pathways but have a NM/NR division, the NM division chiefs were considered as surrogates for a program director to measure Twitter presence (ie, if the division chair had a Twitter handle, it was considered equivalent to program director Twitter presence, even if the division did not offer NM or NR certification). The search was done by two people to improve yield and minimize data deletion.

We then conducted a search cross-matching all potential NM- or NR-related hashtags (Table 1) for each program with a Twitter handle and for program directors with Twitter accounts. These hashtags were selected to broadly encompass the different aspects of this specialty; we also included hashtags used by the Society of Nuclear Medicine and Molecular Imaging (SNMMI) in its tweets. A post by a program with any of these tags between January-December 2019 was considered as positive for NM- or NR-related Twitter activity, irrespective of the number of tweets. Similarly, program director Twitter activity was defined as at least one NM- or NR-related tweet by the program director in the year 2019. Although this search was done for all potential

combinations of hashtags related to NM/NR, #nucmed and #molrad are the only radiology ontology hashtags currently catalogued by the social media analytic tool Symplur [13]. A further subanalysis of individual tweets for the month of December 2019 was done using Symplur and the contents of tweets were categorized as related to radiology education, patient education, departmental promotion, conference talks or lectures, research, and peer networking. Lastly, we conducted a brief survey to assess trends with respect to factors that would

obstruct or facilitate NM/NR physicians' engagement on social media. A SurveyMonkey survey was sent to all email addresses for NM/NR programs available through the Nuclear Medicine Program Directors Association (NMPDA). In total, approximately 45 physicians or program staff were contacted. The survey was anonymized but it allowed the respondents to self-identify as a program director, associate program director, program faculty, or program coordinator. The survey questions are enumerated in [Multimedia Appendix 1](#).

Table 1. List of potential nuclear medicine and nuclear radiology hashtags queried to measure the Twitter presence of academic radiology programs.

Twitter hashtag	Reason selected
#nucmed	Hashtag for nuclear medicine mentioned in the radiology ontology project
#molrad	Hashtag for molecular imaging mentioned in radiology ontology project
#nuclearmedicine	Full expansion of subspecialty
#nuclearradiology	Full expansion of subspecialty
#molecularimaging	Full expansion of subspecialty
#petct	Popular hybrid imaging modality
#petmri	Upcoming hybrid imaging modality
#precisionmedicine	Used by Society of Nuclear Medicine and Molecular Imaging in its tweets to promote the field
#precisionimaging	Used by Society of Nuclear Medicine and Molecular Imaging in its outreach tweets to promote the field
#FOAM #nucmed	Hashtag used to promote Free and Open Access to Medical Education (FOAM) on Twitter, considered synonymous with education
#FOAMrad #nucmed	Hashtag used to promote Free and Open Access Medical Education, considered synonymous with radiology trainee education

Results

For 2019-2020, 39 unique programs were included under NM residency, 15 under NR fellowship, and 34 under the DR/NR category. The 34 programs in the DR/NR category only offered the DR/NR pathway without NR fellowships. Thus, 88 radiology

programs offered training pathways to certification in NM or NR. Of the other 103 radiology residencies, the residents may have only been exposed to NM rotations as part of their radiology residency, without a formal pathway for certification. The results for Twitter engagement by radiology programs and program directors are summarized in [Table 2](#).

Table 2. Summary of Twitter activity related to nuclear medicine or nuclear radiology for radiology programs in 2019-2020.

Activity	NM ^a residency (n=39), n (%)	NR ^b fellowship (n=15), n (%)	Dual DR ^c /NR pathway (n=34), n (%)	Other radiology residencies (n=103), n (%)
Radiology Twitter handle	24 (62)	8 (53)	14 (41)	16 (16)
NM- or NR-specific Twitter handle	3 (8)	0 (0)	0 (0)	0 (0)
Radiology handles tweeting about NM- or NR-related content	13 (33)	4 (27)	4 (9)	5 (5)
Program directors with Twitter handles	3 (8)	4 (27)	0 (0)	0 (0)
Program directors tweeting about NM- or NR-related content	2 (5)	1 (7)	0 (0)	0 (0)

^aNM: nuclear medicine.

^bNR: nuclear radiology.

^cDR: diagnostic radiology.

Out of all programs offering NM/NR training pathways, 46/88 (52%) had radiology accounts on Twitter but only 3/88 (4%) had an exclusive Twitter handle for NM or NR. Of these 88 programs, 21/88 (24%) had at least one tweet related to NM/NR. However, programs offering only DR/NR pathways were less likely to have tweeted about NM- or NR-related content (4/34, 9%) than NM residencies (13/39, 33%) and NR fellowships (4/15, 27%; [Table 1](#)). Of the other 103 radiology residencies,

16/103 (16%) had Twitter handles for radiology, with only 5/103 (5%) tweeting about NM- or NR-related content. The program directors' presence was also low as only 7/88 (8%) NM/NR program directors had Twitter handles. None of the NM division chiefs from the other 103 radiology residencies had an identifiable Twitter handle. Only 3/191 (2%) of program directors actively tweeted about NM/NR ([Table 2](#)).

A content-based subanalysis of NM- and NR-related tweets in December 2019 cross-referenced against hashtags revealed 6 primary tweets by 3 programs for #nucmed, 5 of which were related to NM talks at the annual meeting of the Radiology Society of North America (RSNA), while the remaining tweet was related to department promotion. There were no tweets for the radiology ontology hashtag #molrad but there were 6 tweets/retweets by 6 departments for #molecularimaging, related to RSNA talks and department lectures (n=5) and department promotion (n=1). There were no tweets by academic NM/NR programs with other hashtags such as #precisionmedicine, #precisionimaging, #nuclearmedicine, #FOAMRad #nucmed, and #FOAMed #nucmed. There were also no tweets by NM/NR programs related to radiology education, patient education, research activities, and peer networking.

The brief survey sent to NM/NR faculty and staff via NMPDA contacts further characterized the status of social media use. Overall, one-third of the people contacted (15/45) responded to the survey. The majority of respondents (12/15, 80%) confirmed that they did not have a Twitter handle for their role in their training program; only 3 (20%) had Twitter handles. The majority (11/15, 73%) also confirmed that their programs did not have a unique Twitter handle for their radiology department or NM/NR divisions; only 4 respondents said that their programs had unique Twitter handles. When forced to rank the deterrents to engagement on Twitter, all the listed issues were considered relevant, with no dominant hurdle (Table 3). Despite these perceived hurdles, the 15 respondents thought that the most compelling reason for social media use may be the perceived value of social media for engaging the younger generation of trainees (Table 4).

Table 3. Forced ranking of reasons not to use Twitter.

Reason	Score (1=lowest, 5=highest)
Limited resources (assistance from staff/time to do the work)	3.3
Lack of clarity of value of social media in education	3.6
Lack of expertise among the program director, associate program director, and coordinator	3.4
Another Twitter handle already provides some coverage for this training program	2.6
Negative prior social media experience in a professional setting	2.7

Table 4. Forced ranking of reasons for using Twitter.

Reason	Score (1=lowest, 5=highest)
Perceived value of social media for the younger generation of trainees	4.6
Free marketing	2
Effective way to highlight your training program	3.2
Networking with other programs/organizations	2.2
Follow trends in education in your subspecialty	2.9

Discussion

This study aimed to measure Twitter engagement by academic NM and NR programs in the United States during 2019-2020. Despite 88 programs offering potential pathways to ABNM or ABR NR subspecialty certification, only 3 programs had exclusive handles for the NM division. Although just over half of the 88 programs (n=46, 52%) had a Twitter handle for the broader radiology department, less than one-fourth (n=21, 24%) of all programs tweeted about content related to NM or NR in 2019. The programs offering only DR/NR pathways tweeted about NM- and NR-related content much less than other NM residencies and NR fellowships. Other radiology residencies without NM or NR training pathways also had low Twitter presence (16/103, 16%) and lower Twitter activity related to NM or NR content. Additionally, only 8% (7/88) of NM/NR programs had a DR/NR program director available on Twitter. These findings indicate that there is a substantial missed opportunity for reaching out to or networking with future trainees, a group that has been shown to be open to using social media in their professional development. Although it may not

always be possible—or may even be against existing institutional policies—to allow separate Twitter handles for individual divisions of a department (eg, NM/NR), the overall paucity of Twitter handles for academic NM/NR programs as well as general radiology programs is somewhat remarkable.

The findings of this study are particularly relevant given recent reports of the increasing workforce demand for NM professionals. When juxtaposed against the lack of early exposure and awareness of NM/NR training among medical students and even radiology residents [11,12], it seems NM as a field is not availing itself of a potential method to engage the increasing number of medical students (#medstudents, #medstudentTwitter) and future radiology residents (#futureradres) who are turning to Twitter in an effort to gather information about residency programs [17]. Thus, there is an inherent need for NM/NR academic programs to improve their presence on social media sites such as Twitter and aim for greater online visibility of their own programs. Having more information about NM/NR as a subspecialty and offering engagement opportunities for mentorship, electives, education, and research experiences will not only benefit the current cohort

of trainees but also build the foundation for future development of the subspecialty. Ultimately, both will directly benefit patients by addressing the existing and increasing demand for NM/NR physicians.

A brief survey done after the analysis of the preliminary data attempted to elicit which common issues may be hurdles for the NM/NR academic community in increasing their social media presence. Although only one-third of people contacted (15/45) responded to the survey, the results are helpful to highlight both the challenges and the benefits to social media use. The respondents included program directors, associate program directors, faculty members, and program staff. The respondents confirmed our observation that most programs and program directors do not have unique Twitter handles for their programs. When asked to do a forced ranking of potential reasons to not use Twitter, lack of clarity on the value of social media, lack of expertise, and limited resources to engage on social media were commonly cited by the respondents (Table 3). Although lack of resources is a common issue in many academic operations, the lack of expertise with social media and prior personal negative experiences on social media are also deterrents. Despite these hurdles, the 15 respondents believed that social media could be valuable for engaging the younger generation of trainees (Table 4).

Though our findings demonstrate the limited social media presence of academic NM/NR and DR/NR programs, the general trend of increased use of social media in health care suggests that this specialty may yet find value in increased social media engagement. There are some specific steps that can be taken to make social media use more valuable for the NM/NR community in the future: (1) streamlined consensus use of hashtags, (2) co-opt professional societies to lead the hashtag initiative, and (3) continue to study utilization of social media at the group level in NM/NR in the context of general health care-related use and provide ongoing feedback to our professional community.

Future work may consider each of these value propositions. As part of greater Twitter engagement by the academic NM/NR community, streamlining hashtags related to NM/NR (as summarized in Table 1) and more consistent use of these hashtags may promote aggregation of NM/NR content and more efficiently connect people looking for this information [18-22]. Although #nucmed and #molrad have found a place in the radiology ontology, not all tweets related to NM/NR use these hashtags. This may be because the few hashtags listed in the radiology ontology project were not developed by nuclear radiologists, resulting in a lack of awareness of these hashtags among those in the field of radiology. Further, while the SNMMI (@SNM_MI) often uses #nuclearmedicine and #precisionmedicine in tweets, they have not used #nucmed and #molrad. The American College of Nuclear Medicine and its Nuclear Medicine Resident Organization (@nmroacnm) could consider collaborating with other professional societies with wider Twitter impact, such as the American College of Radiology (ACR), in promoting their specialty. Another option could include reaching out to the Association of Program Directors in Radiology (@theAPDR) to advocate for more

frequent discussion of NM/NR and DR/NR training pathways online and specifically seeking out opportunities to engage medical students and radiology residents early in their training. NM societies such as the SNMMI should also consider leading the community in increasing social media engagement. Just as the RSNA and the ACR focused substantially on social media at their recent annual meetings [20,23], the SNMMI may consider planning sessions focused on harnessing the potential power of social media both during NM society meetings and for ongoing conversations.

This study has several limitations. First, this is a preliminary study that only evaluated engagement on Twitter; other social media networks were not evaluated. We also did not compare the engagement of academic NM with other related specialties such as medical oncology, radiation oncology, and nuclear cardiology, which comprise the top three specialties allied with NM. However, we assessed the lack of engagement against the backdrop of the broad inherent possibility of potential professional interactions at large. We were only able to perform a content-based subanalysis for one month using the free version of Symplur and our search process was manual. We also did not assess the number of followers each program had, or the retweets or social media influence of tweets themselves. Although social media use is higher among younger trainees and medical students [24], we did not have the means to readily correlate the age of program directors, the size of programs, and the number of trainees enrolled as these data are not freely available. These limitations also highlight the general need for more rigorous analytics on social media use in health care. We only considered tweets and retweets by academic NM/NR programs and did not include tweets by people in private radiology practice, scientists, and industry partners who may have higher levels of engagement on Twitter [25]. However, if any of these tweets were retweeted by a particular academic radiology program, it was considered as positive for Twitter presence for that program. The survey did not have a free-text option to capture other responses. Further studies with a broader scope are needed to address each of these limitations. A survey with a larger sample size and a qualitative or free-text component may help us develop a better understanding of barriers to and reasons for social media use in the NM/NR community.

There is currently very little Twitter engagement by the academic nuclear medicine community. This is adequately corroborated by scan data of Twitter use and our attempt to survey faculty in the discipline. Although there are identifiable obstacles, the responses by NM faculty and staff as well as the general trend of increased social media use among medical students substantially support the perceived value in increasing social media engagement in imaging specialties. A more in-depth investigation in the future may further help us understand the barriers and benefits of social media use, and assess the impact of increased use on trainee recruitment and perceptions. Additionally, the value proposition of streamlining and growing social media engagement with targeted hashtags may be considered to promote the presence of both diagnostic and therapeutic aspects of this subspecialty for practitioners, trainees, and the public.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questions.

[[DOCX File, 22 KB - formative_v5i11e24448_app1.docx](#)]

References

1. Social Media Fact Sheet. Pew Research Center. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/> [accessed 2020-07-01]
2. Patel SS, Hawkins CM, Rawson JV, Hoang JK. Professional Social Networking in Radiology: Who Is There and What Are They Doing? *Acad Radiol* 2017 May;24(5):574-579. [doi: [10.1016/j.acra.2016.09.026](https://doi.org/10.1016/j.acra.2016.09.026)] [Medline: [28153576](https://pubmed.ncbi.nlm.nih.gov/28153576/)]
3. Ranschaert ER, Van Ooijen PMA, McGinty GB, Parizel PM. Radiologists' Usage of Social Media: Results of the RANSOM Survey. *J Digit Imaging* 2016 Feb 3;29(4):443-449. [doi: [10.1007/s10278-016-9865-1](https://doi.org/10.1007/s10278-016-9865-1)]
4. Diug B, Kendal E, Ilic D. Evaluating the use of twitter as a tool to increase engagement in medical education. *Educ Health (Abingdon)* 2016;29(3):223-230 [FREE Full text] [doi: [10.4103/1357-6283.204216](https://doi.org/10.4103/1357-6283.204216)] [Medline: [28406107](https://pubmed.ncbi.nlm.nih.gov/28406107/)]
5. Kohli MD, Daye D, Towbin AJ, Kotsenas AL, Heilbrun ME. Social Media Tools for Department and Practice Communication and Branding in the Digital Age. *Radiographics* 2018 Oct;38(6):1773-1785. [doi: [10.1148/rg.2018180090](https://doi.org/10.1148/rg.2018180090)] [Medline: [30303796](https://pubmed.ncbi.nlm.nih.gov/30303796/)]
6. Miles R, Patel AK. The Radiology Twitterverse: A Starter's Guide to Utilization and Success. *J Am Coll Radiol* 2019 Sep;16(9 Pt A):1225-1231. [doi: [10.1016/j.jacr.2019.03.014](https://doi.org/10.1016/j.jacr.2019.03.014)] [Medline: [31092350](https://pubmed.ncbi.nlm.nih.gov/31092350/)]
7. Heitkamp DE, Cooke EA, Deitte LA, Hunt KN, Milburn JM, Patel AK, et al. Radiology Program Directors Should Have an Active Presence on Twitter. *J Am Coll Radiol* 2020 Feb;17(2):293-295. [doi: [10.1016/j.jacr.2019.10.014](https://doi.org/10.1016/j.jacr.2019.10.014)] [Medline: [31843345](https://pubmed.ncbi.nlm.nih.gov/31843345/)]
8. Ranginwala S, Towbin AJ. The Power of Promotion: Using Social Media to Promote a Radiology Department. *Acad Radiol* 2017 Apr;24(4):488-496. [doi: [10.1016/j.acra.2016.11.011](https://doi.org/10.1016/j.acra.2016.11.011)] [Medline: [31203949](https://pubmed.ncbi.nlm.nih.gov/31203949/)]
9. Redmond CE, Gibney B, Nicolaou S, Forster BB. Recruiting the Next Generation of Radiologists: The Important Role of Social Media. *Acad Radiol* 2020 Sep;27(9):1335. [doi: [10.1016/j.acra.2019.11.013](https://doi.org/10.1016/j.acra.2019.11.013)] [Medline: [31866108](https://pubmed.ncbi.nlm.nih.gov/31866108/)]
10. Arevalo-Perez J, Paris M, Graham MM, Osborne JR. A Perspective of the Future of Nuclear Medicine Training and Certification. *Semin Nucl Med* 2016 Jan;46(1):88-96 [FREE Full text] [doi: [10.1053/j.semnuclmed.2015.10.003](https://doi.org/10.1053/j.semnuclmed.2015.10.003)] [Medline: [26687859](https://pubmed.ncbi.nlm.nih.gov/26687859/)]
11. Ng TS, An BP, Cho SY, Hyun H. US Trainee and Faculty Perspectives on Exposure to Nuclear Medicine/Molecular Imaging During Medical School. *Curr Probl Diagn Radiol* 2021 Sep;50(5):585-591. [doi: [10.1067/j.cpradiol.2020.05.013](https://doi.org/10.1067/j.cpradiol.2020.05.013)] [Medline: [32653243](https://pubmed.ncbi.nlm.nih.gov/32653243/)]
12. Ruddell JH, Eltorai AE, Tang OY, Suskin JA, Dibble EH, Oates ME, et al. The Current State of Nuclear Medicine and Nuclear Radiology: Workforce Trends, Training Pathways, and Training Program Websites. *Acad Radiol* 2020 Dec;27(12):1751-1759. [doi: [10.1016/j.acra.2019.09.026](https://doi.org/10.1016/j.acra.2019.09.026)] [Medline: [31759795](https://pubmed.ncbi.nlm.nih.gov/31759795/)]
13. Hawkins CM. Radiology's Social Media Hashtag Ontology: Codifying Online Data. *J Am Coll Radiol* 2016 Jan;13(1):111-113. [doi: [10.1016/j.jacr.2015.08.028](https://doi.org/10.1016/j.jacr.2015.08.028)] [Medline: [26499165](https://pubmed.ncbi.nlm.nih.gov/26499165/)]
14. Training Requirements for the ABNM Certifying Examination. The American Board of Nuclear Medicine. URL: <https://www.abnm.org/index.php/exam/training-requirements/> [accessed 2019-10-30]
15. Nuclear Radiology: Requirements and Registration. American Board of Radiology. 2019. URL: <https://www.theabr.org/diagnostic-radiology/subspecialties/nuclear-radiology/requirements-registration> [accessed 2019-10-30]
16. Nuclear Medicine Programs 2019-2020. ACGME. URL: <https://apps.acgme.org/ads/Public/Reports/Report/1> [accessed 2019-10-30]
17. Nguyen JK, Shah N, Heitkamp DE, Gupta Y. COVID-19 and the Radiology Match: A Residency Program's Survival Guide to the Virtual Interview Season. *Acad Radiol* 2020 Sep;27(9):1294-1297 [FREE Full text] [doi: [10.1016/j.acra.2020.06.023](https://doi.org/10.1016/j.acra.2020.06.023)] [Medline: [32660754](https://pubmed.ncbi.nlm.nih.gov/32660754/)]
18. Mahmari JJ. Fresh initiatives reflected at ASNC2012. *J Nucl Cardiol* 2012 Aug;19(4):870-871. [doi: [10.1007/s12350-012-9595-9](https://doi.org/10.1007/s12350-012-9595-9)] [Medline: [22753074](https://pubmed.ncbi.nlm.nih.gov/22753074/)]
19. Bhargava P. Using Twitter to Amplify Your Presence During Radiology Meetings. *Curr Probl Diagn Radiol* 2017 Jul;46(4):261-262. [doi: [10.1067/j.cpradiol.2017.05.010](https://doi.org/10.1067/j.cpradiol.2017.05.010)] [Medline: [28641820](https://pubmed.ncbi.nlm.nih.gov/28641820/)]
20. Hawkins CM, Duszak R, Rawson JV. Social media in radiology: early trends in Twitter microblogging at radiology's largest international meeting. *J Am Coll Radiol* 2014 Apr;11(4):387-390. [doi: [10.1016/j.jacr.2013.07.015](https://doi.org/10.1016/j.jacr.2013.07.015)] [Medline: [24139963](https://pubmed.ncbi.nlm.nih.gov/24139963/)]
21. Kalia V, Ortiz DA, Patel AK, Moriarity AK, Canon CL, Duszak R. Leveraging Twitter to Maximize the Radiology Meeting Experience. *J Am Coll Radiol* 2018 Jan;15(1 Pt B):177-183. [doi: [10.1016/j.jacr.2017.10.022](https://doi.org/10.1016/j.jacr.2017.10.022)] [Medline: [29162419](https://pubmed.ncbi.nlm.nih.gov/29162419/)]

22. Sheikh S, Patel MV, Song Y, Navuluri R, Zangan S, Ahmed O. Social Media Growth at Annual Medical Society Meetings: A Comparative Analysis of Diagnostic and Interventional Radiology to Other Medical Specialties. *Curr Probl Diagn Radiol* 2021 Sep;50(5):592-598. [doi: [10.1067/j.cpradiol.2020.06.001](https://doi.org/10.1067/j.cpradiol.2020.06.001)] [Medline: [32654834](https://pubmed.ncbi.nlm.nih.gov/32654834/)]
23. Rosenkrantz AB, Hawkins CM. Use of Twitter Polls to Determine Public Opinion Regarding Content Presented at a Major National Specialty Society Meeting. *J Am Coll Radiol* 2017 Feb;14(2):177-182. [doi: [10.1016/j.jacr.2016.07.024](https://doi.org/10.1016/j.jacr.2016.07.024)] [Medline: [27687748](https://pubmed.ncbi.nlm.nih.gov/27687748/)]
24. Koontz NA, Kamer AP, Dodson SC, Capps AE, Tomblinson CM, Brown BP, et al. Social Media Utilization at an Academic Radiology Practice. *Acad Radiol* 2018 Jan;25(1):111-117. [doi: [10.1016/j.acra.2017.08.012](https://doi.org/10.1016/j.acra.2017.08.012)] [Medline: [29056400](https://pubmed.ncbi.nlm.nih.gov/29056400/)]
25. Glover M, Choy G, Boland GW, Saini S, Prabhakar AM. Radiology and social media: are private practice radiology groups more social than academic radiology departments? *J Am Coll Radiol* 2015 May;12(5):513-518. [doi: [10.1016/j.jacr.2014.11.005](https://doi.org/10.1016/j.jacr.2014.11.005)] [Medline: [25707884](https://pubmed.ncbi.nlm.nih.gov/25707884/)]

Abbreviations

ABR: American Board of Radiology
ACR: American College of Radiology
DR: diagnostic radiology
NM: nuclear medicine
NMPDA: Nuclear Medicine Program Directors Association
NR: nuclear radiology
RSNA: Radiology Society of North America
SNMMI: Society of Nuclear Medicine and Molecular Imaging

Edited by G Eysenbach; submitted 22.09.20; peer-reviewed by S Kotcherlakota, S Malecki; comments to author 21.10.20; revised version received 16.03.21; accepted 20.09.21; published 08.11.21.

Please cite as:

Panda A, Sharma A, Dundar A, Packard A, Aase L, Kotsenas A, Kendi AT
Twitter Use by Academic Nuclear Medicine Programs: Pilot Content Analysis Study
JMIR Form Res 2021;5(11):e24448
URL: <https://formative.jmir.org/2021/11/e24448>
doi: [10.2196/24448](https://doi.org/10.2196/24448)
PMID: [34747708](https://pubmed.ncbi.nlm.nih.gov/34747708/)

©Ananya Panda, Akash Sharma, Ayca Dundar, Ann Packard, Lee Aase, Amy Kotsenas, Ayse Tuba Kendi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 08.11.2021. 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

Googling for Suicide—Content and Quality Analysis of Suicide-Related Websites: Thematic Analysis

Wen Chen¹, MSc; Andrea Boggero², PsyD; Giovanni Del Puente³, MD; Martina Olcese², PsyD; Davide Prestia⁴, PhD, MD; Haitham Jahrami^{5,6}, PhD; Nasr Chalhaf⁷, PhD; Noomen Guelmami⁸, PhD; Fairouz Azaiez⁷, PhD; Nicola Luigi Bragazzi⁹, MPH, MD, PhD

¹Department of Psychiatry, Xiamen Xianyue Hospital, Xiamen, China

²Department of Educational Science, University of Genoa, Genoa, Italy

³Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics and Maternal/Child Sciences, University of Genoa, Genoa, Italy

⁴Department of Psychiatry, Istituto di ricovero e cura a carattere scientifico Ospedale Policlinico San Martino, Genoa, Italy

⁵College of Medicine and Medical Sciences, Arabian Gulf University, Manama, Bahrain

⁶Ministry of Health, Manama, Bahrain

⁷Higher Institute of Sport and Physical Education of Sfax, University of Sfax, Sfax, Tunisia

⁸Higher Institute of Sport and Physical Education of Kef, University of Jendouba, Jendouba, Tunisia

⁹Laboratory for Industrial and Applied Mathematics, Department of Mathematics and Statistics, York University, Toronto, ON, Canada

Corresponding Author:

Nicola Luigi Bragazzi, MPH, MD, PhD

Laboratory for Industrial and Applied Mathematics

Department of Mathematics and Statistics

York University

4700 Keele St

Toronto, ON, M3J 1P3

Canada

Phone: 1 416 736 2100

Email: bragazzi@yorku.ca

Abstract

Background: Suicide represents a public health concern, imposing a dramatic burden. Prosuicide websites are “virtual pathways” facilitating a rise in suicidal behaviors, especially among socially isolated, susceptible individuals.

Objective: The aim of this study is to characterize suicide-related webpages in the Italian language.

Methods: The first 5 most commonly used search engines in Italy (ie, Bing, Virgilio, Yahoo, Google, and Libero) were mined using the term “suicidio” (Italian for suicide). For each search, the first 100 webpages were considered. Websites resulting from each search were collected and duplicates deleted so that unique webpages could be analyzed and rated with the HONcode instrument

Results: A total of 65 webpages were included: 12.5% (8/64) were antisuicide and 6.3% (4/64) explicitly prosuicide. The majority of the included websites had a mixed or neutral attitude toward suicide (52/64, 81.2%) and had informative content and purpose (39/64, 60.9%). Most webpages targeted adolescents as an age group (38/64, 59.4%), contained a reference to other psychiatric disorders or comorbidities (42/64, 65.6%), included medical/professional supervision or guidance (45/64, 70.3%), lacked figures or pictures related to suicide (41/64, 64.1%), and did not contain any access restraint (62/64, 96.9%). The major shortcoming to this study is the small sample size of webpages analyzed and the search limited to the keyword “suicide.”

Conclusions: Specialized mental health professionals should try to improve their presence online by providing high-quality material.

(*JMIR Form Res* 2021;5(11):e29146) doi:[10.2196/29146](https://doi.org/10.2196/29146)

KEYWORDS

suicide; internet; world wide web; content analysis; HONcode; mental health; webpage; health information; eHealth

Introduction

Suicide represents a public health concern, imposing a dramatic burden both in epidemiological and societal terms [1]. According to the latest available 2021 report of the World Health Organization (WHO) in 2019, Italy had a crude rate of 6.7 (95% CI 6.0-8.5) suicides per 100,000 people, with a male:female ratio of 2.9 [2].

Within the current eHealth era, in which the new information and communication technologies are being increasingly exploited to obtain health-related information [3], the internet can be characterized as having a Yin-Yang, paradoxical nature [4]. On the one hand, there exists a statistically significant correlation between a pathological use of the web and suicidality (either in terms of suicidal ideation or nonsuicidal self-injury). Prosuicide websites and online suicide pacts are “virtual pathways” that facilitate the emergence of suicidal behaviors, especially among socially isolated and susceptible individuals. On the other hand, the internet can serve as an effective tool for suicide monitoring and tracking [3,5] and can play a major role in suicide prevention [6,7], particularly for those vulnerable individuals who would be otherwise difficult to reach with conventional approaches [8,9].

Understanding the determinants of suicide-related web searches is, therefore, of crucial importance [10]. According to Wong and collaborators [11], the categories of the most visited suicide-related webpages can be classified into entertainment (30.13%), scientific information (18.31%), and community resources (14.53%). Among the 1314 accessed webpages in their study, prosuicide websites represented only a small fraction (0.15%). The most used search terms were “committing suicide with a gas oven,” “hairless goat,” “pictures of murder by strangulation,” and “photo of a severe burn.”

Suicide-related digital behavior has been assessed in different countries and different languages [12-15]. However, to the best of our knowledge, there is a dearth of information concerning suicide-related webpages in the Italian language. As such, the aim of this study is to address this gap in knowledge.

Methods

Suicide-Related Website Selection and Inclusion

The first 5 most commonly used search engines in Italy (namely, Bing, Virgilio, Yahoo, Google, and Libero) were mined with the search term “suicidio” (Italian for suicide). For each search, the first 100 webpages were considered. Websites resulting from each search were collected and duplicates deleted so that webpages could be analyzed and rated by 2 authors independently (MO and DP).

Quality assessment of the included webpages was performed with the HONcode instrument, which comprises the following sections: authoritative (if the qualifications of the authors are mentioned), complementarity (if the site aims at supporting but not replacing or substituting the physician-patient relationship), privacy (if visitors' data are protected and respected), attribution (if sources of published information in the medical and health-related pages are explicitly mentioned and referred to),

justifiability (if claims related to potential benefits or performance are evidence based), transparency (if the presentation of the online material is accessible, with accurate email contacts), financial disclosure (if eventual funding sources are identified and mentioned), and advertising policy (if advertising and editorial materials are separate and identifiable) [16].

The agreement between the 2 authors was assessed with the Cohen κ statistic.

Furthermore, the 2 authors filled in an ad hoc predesigned Excel spreadsheet (Microsoft Corp) containing questions concerning the attitude of the webpage toward suicide (prosuicide, antisuicide, mixed or neutral), content and purpose of the webpage (supporting or emotional, informative, preventive), supervision or guidance (if the webpage was user-generated or handled by a professional figure, such as a psychiatrist or a psychologist), reference to age groups (adolescents or older people) targeted or referred to by the webpage (this simple dichotomic stratification was applied based on the availability of information provided and considering that the majority of internet users consist of adolescents), reference to other psychiatric disorders or comorbidities, the presence of figures or pictures related to suicide, and the presence of access restraint.

Moreover, for each included webpage, the number of comments' likes, and shares, if any, was recorded.

Statistical Analysis

Before commencing any statistical handling and processing of the data, figures were visually inspected for potential outliers. The normality of data distribution was checked using the D'Agostino-Pearson omnibus test. Univariate analysis (chi-squared test and the chi-squared test for trend, Fisher's exact test, *t* test, and analysis of variance or their nonparametric versions in case of violation of normal distribution) was performed for the variables under study. Multivariate linear regression analysis was performed to shed light on the determinants of the HONcode score.

All statistical analyses were performed with the commercial SPSS for Windows, version 24.0 (IBM Corp). Figures with *P* values less than .05 were considered statistically significant.

Results

Sixty-four unique websites were retrieved and analyzed. Concerning the attitude toward suicide, 4 webpages were in favor of suicide (6.3%), whereas 8 sites were against suicide (12.5%). The remaining webpages (52/64, 81.2%) were neutral toward suicide. The overrepresentation of neutral pages in tone was statistically significant ($\chi^2_2=30.55$; contingency coefficient=0.44; $P<.001$; χ^2_1 for trend=28.40; $P<.001$). In terms of content and purpose, 6 webpages were supporting or emotional (9.4%), 39 informative (60.9%), and 19 preventive (29.7%), with both emotional and preventive pages being underrepresented ($\chi^2_2=13.83$, contingency coefficient=0.31, $P=.001$; χ^2_1 for trend=13.67, $P<.001$). Concerning the

supervision or guidance of the websites, 45 (70.3% of the entire sample) were handled by a professional figure (a medical doctor or a psychologist, or an allied health professional). Consequently, the medical supervision or guidance of the webpages was predominant ($P=.03$). Thirty-eight webpages were targeted to or contained references to adolescents as an age group particularly vulnerable to suicide (59.4%; 26/64 websites referred to older people, 40.6%; $P=.37$). Forty-two (65.6%) webpages made references to other psychiatric disorders or comorbidities, such as depression (65.6%) or drug use (26.6%), while twenty-two (34.4%) websites did not. Webpages mentioning comorbidities were slightly overrepresented ($P=.07$). Twenty-three webpages (35.9%) versus forty-one (64.1%) contained figures or pictures related to suicide (not statistically significant, $P=.15$). Sixty-two webpages (96.9% of the sample) had no access restraint ($P<.001$).

The average HONcode score was 5.81 (SD 2.70, median 7), whereas the average overall number of likes, comments, and shares was 14,261.56 (SD 60,065.23, median 15.00). The correlation between HONcode and the overall number of likes, comments, and shares was 0.16 (95% CI -0.33 to 0.58; $P=.53$). Further details are reported in [Table 1](#).

Webpages were generally compliant with the HONcode principles (range 68.8%-100.0%; authoritative principle: 44/64; principle of transparency: 64/64) except for the principle of justifiability (35/64, 54.7%, $P=.72$) and financial disclosure (40/64, 62.5%, $P=.21$), as shown in [Table 2](#).

At the univariate analysis ([Table 3](#)), the HONcode score was higher in mixed or neutral webpages ($P<.001$), in websites with a preventive purpose ($P=.02$), with professional or medical supervision or guidance ($P<.001$), with references to other psychiatric disorders or comorbidities ($P=.003$), and without access restraint ($P=.03$). No statistically significant differences could be detected in terms of references to age groups (adolescents versus older people, $P=0.81$) or the presence of figures or pictures ($P=.41$).

Furthermore, webpages with medical supervision or guidance had more an informative and preventive content and purpose than did supporting or emotional ones ($\chi^2_2=9.19$, $P=.01$; χ^2_1 for trend=3.19, $P=.07$). Those webpages with professional supervision or guidance also had a more positive or mixed/neutral attitude toward suicide, and this was statistically significant ($P=.002$). In contrast, no differences were found in terms of references to other psychiatric disorders or comorbidities, references to age groups, access restraint, and the presence of figures or pictures related to suicide.

In the multivariate linear regression analysis ([Table 4](#)), the predictors of the HONcode score were professional/medical or supervision/guidance of the webpages (regression coefficient=3.94; $P<.001$) and references to psychiatric disorders as comorbidities (regression coefficient=1.46; $P=.003$).

Meanwhile, no differences were found in terms of likes, comments, or shares either in the univariate or the multivariate analyses.

Table 1. Characteristics of the included websites (N=64).

Parameter	Value
Attitude toward suicide, n (%)	
Prosuicide	4 (6.3)
Antisuicide	8 (12.5)
Mixed/neutral	52 (81.2)
Content and purpose, n (%)	
Supporting/emotional	6 (9.4)
Informative	39 (60.9)
Preventive	19 (29.7)
Supervision/guidance, n (%)	
Religious figure	4 (6.3)
Psychiatrist	11 (17.2)
Other kinds of physicians	3 (4.7)
Psychologist	25 (39.1)
Pharmacist	1 (1.6)
Other allied health professionals	5 (7.8)
Lawyer	1 (1.6)
User-generated webpage	11 (17.2)
Volunteer organization/charity	3 (4.7)
References to age groups, n (%)	
Adolescents	38 (59.4)
Older people	26 (40.6)
Reference to other psychiatric disorders, n (%)	
Depression	42 (65.6)
Anxiety	11 (17.2)
Panic disorder	2 (3.1)
Bipolar syndrome	9 (14.1)
Personality disorder	8 (12.5)
Posttraumatic stress disorder	3 (4.7)
Paranoia	2 (3.1)
Schizophrenia	12 (18.8)
Drug use	17 (26.6)
Alcoholism	6 (9.4)
Eating and weight disorders	1 (1.6)
Sleep disorders	1 (1.6)
Reference	42 (65.6)
No reference	22 (34.4)
Figures/pictures, n (%)	
Yes	23 (35.9)
No	41 (64.1)
Access restraint, n (%)	
Yes	2 (3.1)
No	62 (96.9)

Parameter	Value
HONcode score, mean (SD)	5.81 (2.7)
Comments, n	585
Shares, n	98
Likes, n	256,025

Table 2. Assessment of the suicide-related websites according to the HONcode principles (N=64).

HONcode principle	Yes	No	P value
1. Authoritative	44 (68.8)	20 (31.2)	.047
2. Complementarily	48 (75.0)	16 (25.0)	.006
3. Privacy	48 (75.0)	16 (25.0)	.006
4. Attribution	45 (70.3)	19 (29.7)	.03
5. Justifiability	35 (54.7)	29 (45.3)	.72
6. Transparency	64 (100.0)	0 (0.0)	<.001
7. Financial disclosure	40 (62.5)	24 (37.5)	.21
8. Advertising policy	50 (78.1)	14 (21.9)	.002

Table 3. Univariate analysis of the studied webpages according to the HONcode score.

Parameter	HONcode score, mean (SD)	P value
Attitude toward suicide		<.001
Prosuicide	0.75 (0.50)	
Antisuicide	3.50 (2.93)	
Mixed/neutral	6.55 (2.07)	
Content and purpose		.02
Supporting/emotional	2.17 (2.93)	
Informative	6.08 (2.45)	
Preventive	6.42 (2.22)	
Supervision/guidance		<.001
User-generated webpage	2.63 (2.27)	
Professional figures	7.16 (1.38)	
References to age groups		.81
Adolescents	6.05 (2.36)	
Older people	5.46 (3.09)	
Reference to other psychiatric disorders		.003
No reference	4.28 (3.16)	
Reference	6.41 (2.22)	
Figures/pictures		.41
Yes	6.09 (2.83)	
No	5.66 (2.60)	
Access restraint		.03
Yes	100 (0.00)	
No	5.97 (2.57)	

Table 4. Multivariate linear regression analysis of the studied webpages according to the HONcode score.

Independent variables	Value				
	Coefficient	SE	r_{partial}	t	P
Constant	1.11	N/A ^a	N/A	N/A	N/A
Content and purpose	0.27	0.20	0.17	1.30	.20
Reference to age groups	0.18	0.44	0.06	0.42	.78
Figures/pictures	0.40	0.45	0.12	0.89	.38
Supervision/guidance	3.94	0.47	0.75	8.30	<.001
Access restraint	-1.99	1.42	-0.19	-1.40	.17
References to other psychiatric disorders	1.46	0.47	0.39	3.12	.003
Attitude toward suicide	0.52	0.57	0.12	0.91	.37

^aN/A: not applicable.

Discussion

Main Findings and Comparison With the Literature

To the best of our knowledge, this is the first study to assess the quality and content of suicide-related webpages in the Italian language.

We found that the percentage of explicitly prosuicide websites was greater than that reported by Wong et al [11] but less than that reported by Chen et al (16.3%), who analyzed 375 linked webpages in Chinese [12]. According to the authors, 41.3% of the included webpages were antisuicide, and the majority of the prosuicide sites were user-generated (96.7%). Interestingly, searches using the search term “ways to kill yourself” (31.7%) and “painless suicide” (28.3%) generated much larger numbers of potentially harmful webpages than did the keyword “suicide” (4.3%).

A much larger portion of prosuicide websites (42%) was found by Sakarya et al [17], who analyzed 100 webpages. Thirteen percent of the studied sample was found to have content or material that may be considered protective against suicidal behavior, a result in line with our findings. However the authors found that “protective” websites did not have any kind of medical supervision or guidance from specialized mental health professionals. In contrast to this, we managed to find a statistically significant association between these 2 variables.

Til and Niederkrotenthaler [18] performed a cross-cultural study by comparing the suicide-related web searches in Austria and the USA. They found that in both countries, protective outweighed harmful website characteristics by approximately 2:1, a finding in line with our result. Authors were able to confirm the observation of Wong et al: websites retrieved with method-related search terms (eg, how to hang yourself) contained more harmful information (USA: $P<.001$; Austria: $P<.05$) and fewer protective characteristics (USA: $P<.001$; Austria: $P<.001$) compared to the term “suicide” [11]. Conversely, help-related search terms (eg, suicidal thoughts) yielded more webpages with positive and protective characteristics (USA: $P=.07$; Austria: $P<.01$). Interestingly, authors found that websites retrieved with the US search engines generally had more protective characteristics ($P<.001$) than did

those retrieved with Austrian search engines. Moreover, resources with dangerous or harmful characteristics were better ranked than those with positive or protective characteristics (USA: $P<.01$; Austria: $P<.05$).

In another study, Thornton and colleagues [13] found that among Google searches retrieved related to suicide, a high portion of them were irrelevant webpages ($n=136$, approximately 26% of the entire sample). Of the 329 relevant websites, the majority were suicide preventive (about 68%); however, a considerable proportion of sites expressed mixed/ambiguous (22%) or neutral (8%) attitudes toward suicide, and 1% were explicitly prosuicide, figures which were slightly different from ours.

These inconsistent literature findings could be explained by the cultural differences across countries in which suicide-related web searches are performed, as shown by the research of Till and Niederkrotenthaler [15], as well as policies concerning the web. Overall, the combination of societal data and online behavior monitoring provide the best indication of risks [19].

The presence of antisuicide online material is of crucial importance for 2 major reasons: (1) according to a recent review of the literature [6], individuals at risk for suicide are probably more prone to look for suicide-related information online, and searching online information related to suicide is a proxy of suicidal ideation; and (2) according to a randomized controlled trial, education professional suicide prevention websites appear to increase and improve suicide prevention-related knowledge, especially among vulnerable and socially isolated individuals [18]. Therefore, exposure to high-quality curated websites may be associated with a reduction in suicidal thoughts and actions. Recent research has shown that the proactive suicide prevention online model is robust for identifying people that are at risk of suicide [20].

Strengths and Limitations of the Study

Despite some strengths, including being the first study to systematically assess the quality of suicide-related webpages in the Italian language, our investigation suffers from some limitations that should be properly acknowledged. The major shortcoming is the small sample size of webpages analyzed and the search limited to the keyword “suicide.” Future research should also explore other words indirectly related to suicide as

well as other terms related to the ways suicide can be committed (ie, method-related search terms) and explore whether the semantic differential has an impact on suicide-related webpages in terms of content and quality. Moreover, caution should be used when generalizing the present findings to other countries, as our results could be affected by country-specific cultural issues and may not reflect the nature of other settings. Further, although we assessed the reference to other psychiatric disorders or comorbidities in the webpages included in our study, we were not able to establish a link between mental illness and suicide in terms of browsing websites. This warrants dedicated, ad hoc studies specifically exploring this topic. Another aspect that should be investigated concerns the size or burden of the psychological effect imposed by browsing suicide-related websites.

Conclusions

Sixty-four webpages were included in our analysis: 12.5% (8/64) were antisuicide and 6.3% (4/64) explicitly pro-suicide. The majority of the included websites had a mixed or neutral attitude toward suicide (52/64, 81.2%) and had an informative content and purpose (39/64, 60.9%). Most webpages targeted adolescents as an age group (38/64, 59.4%), contained a reference to other psychiatric disorders or comorbidities (42/64, 65.6%), had medical or professional supervision or guidance (45/64, 70.3%), lacked figures or pictures related to suicide (41/64, 64.1%), and did not use any access restraint (62/64, 96.9%).

Specialized mental health professionals should take into account these findings, make an effort to improve their presence online, and provide high-quality material. However, given the aforementioned limitations, further research in the field is warranted.

Conflicts of Interest

None declared.

References

1. Sinyor M, Tse R, Pirkis J. Global trends in suicide epidemiology. *Curr Opin Psychiatry* 2017 Jan;30(1):1-6. [doi: [10.1097/YCO.0000000000000296](https://doi.org/10.1097/YCO.0000000000000296)] [Medline: [27845946](https://pubmed.ncbi.nlm.nih.gov/27845946/)]
2. Suicide in the world. Global health estimates. World Health Organization. Geneva, Switzerland; 2019. URL: <https://apps.who.int/iris/bitstream/handle/10665/326948/WHO-MSD-MER-19.3-eng.pdf> [accessed 2021-11-02]
3. Bragazzi NL. A Google Trends-based approach for monitoring NSSI. *Psychol Res Behav Manag* 2013;7(6):1-8 [FREE Full text] [doi: [10.2147/PRBM.S44084](https://doi.org/10.2147/PRBM.S44084)] [Medline: [24376364](https://pubmed.ncbi.nlm.nih.gov/24376364/)]
4. Durkee T, Hadlaczky G, Westerlund M, Carli V. Internet pathways in suicidality: a review of the evidence. *Int J Environ Res Public Health* 2011 Oct;8(10):3938-3952 [FREE Full text] [doi: [10.3390/ijerph8103938](https://doi.org/10.3390/ijerph8103938)] [Medline: [22073021](https://pubmed.ncbi.nlm.nih.gov/22073021/)]
5. Solano P, Ustulin M, Pizzorno E, Vichi M, Pompili M, Serafini G, et al. A Google-based approach for monitoring suicide risk. *Psychiatry Res* 2016 Dec 30;246:581-586. [doi: [10.1016/j.psychres.2016.10.030](https://doi.org/10.1016/j.psychres.2016.10.030)] [Medline: [27837725](https://pubmed.ncbi.nlm.nih.gov/27837725/)]
6. Mok K, Jorm AF, Pirkis J. Suicide-related Internet use: A review. *Aust N Z J Psychiatry* 2015 Aug;49(8):697-705. [doi: [10.1177/0004867415569797](https://doi.org/10.1177/0004867415569797)] [Medline: [25698810](https://pubmed.ncbi.nlm.nih.gov/25698810/)]
7. Padmanathan P, Biddle L, Carroll R, Derges J, Potokar J, Gunnell D. Suicide and self-harm related internet use. *Crisis* 2018 Nov;39(6):469-478 [FREE Full text] [doi: [10.1027/0227-5910/a000522](https://doi.org/10.1027/0227-5910/a000522)] [Medline: [29848080](https://pubmed.ncbi.nlm.nih.gov/29848080/)]
8. Jacob N, Scourfield J, Evans R. Suicide prevention via the Internet: a descriptive review. *Crisis* 2014 Jan 01;35(4):261-267. [doi: [10.1027/0227-5910/a000254](https://doi.org/10.1027/0227-5910/a000254)] [Medline: [24984890](https://pubmed.ncbi.nlm.nih.gov/24984890/)]
9. Christensen H, Batterham PJ, O'Dea B. E-health interventions for suicide prevention. *Int J Environ Res Public Health* 2014 Aug;11(8):8193-8212 [FREE Full text] [doi: [10.3390/ijerph110808193](https://doi.org/10.3390/ijerph110808193)] [Medline: [25119698](https://pubmed.ncbi.nlm.nih.gov/25119698/)]
10. Marchant A, Hawton K, Stewart A, Montgomery P, Singaravelu V, Lloyd K, et al. A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. *PLoS ONE* 2017 Aug 16;12(8):e0181722. [doi: [10.1371/journal.pone.0181722](https://doi.org/10.1371/journal.pone.0181722)]
11. Wong PW, Fu K, Yau RS, Ma HH, Law Y, Chang S, et al. Accessing suicide-related information on the internet: a retrospective observational study of search behavior. *J Med Internet Res* 2013 Jan 11;15(1):e3 [FREE Full text] [doi: [10.2196/jmir.2181](https://doi.org/10.2196/jmir.2181)] [Medline: [23305632](https://pubmed.ncbi.nlm.nih.gov/23305632/)]
12. Chen Y, Hung GC, Cheng Q, Tsai C, Wu KC. Searching for suicide-related information on Chinese websites. *Psychiatry Res* 2017 Dec;258:506-510. [doi: [10.1016/j.psychres.2017.08.087](https://doi.org/10.1016/j.psychres.2017.08.087)] [Medline: [28886904](https://pubmed.ncbi.nlm.nih.gov/28886904/)]
13. Thornton L, Handley T, Kay-Lambkin F, Baker A. Is a person thinking about suicide likely to find help on the internet? An evaluation of Google search results. *Suicide Life Threat Behav* 2017 Feb;47(1):48-53. [doi: [10.1111/sltb.12261](https://doi.org/10.1111/sltb.12261)] [Medline: [27111845](https://pubmed.ncbi.nlm.nih.gov/27111845/)]
14. Singaravelu V, Stewart A, Adams J, Simkin S, Hawton K. Information-seeking on the internet. *Crisis* 2015;36(3):211-219. [doi: [10.1027/0227-5910/a000307](https://doi.org/10.1027/0227-5910/a000307)] [Medline: [26088826](https://pubmed.ncbi.nlm.nih.gov/26088826/)]
15. Till B, Niederkrotenthaler T. Surfing for suicide methods and help. *J. Clin. Psychiatry* 2014 Jul 08;75(08):886-892. [doi: [10.4088/jcp.13m08861](https://doi.org/10.4088/jcp.13m08861)]
16. HONcode: Principles. HealthOnNet. 1995. URL: <https://www.healthonnet.org/HONcode/Conduct.html> [accessed 2021-11-02]

17. Sakarya D, Güneş C, Sakarya A. Googling suicide: evaluation of websites according to the content associated with suicide. *Turk Psikiyatri Derg* 2013;24(1):44-48 [FREE Full text] [Medline: [23446539](#)]
18. Till B, Tran US, Voracek M, Niederkrotenthaler T. Beneficial and harmful effects of educative suicide prevention websites: randomised controlled trial exploring Papageno v. Werther effects. *Br J Psychiatry* 2017 Aug;211(2):109-115. [doi: [10.1192/bjp.bp.115.177394](#)] [Medline: [28522433](#)]
19. Barros JM, Melia R, Francis K, Bogue J, O'Sullivan M, Young K, et al. The Validity of Google Trends Search Volumes for Behavioral Forecasting of National Suicide Rates in Ireland. *Int J Environ Res Public Health* 2019 Sep 02;16(17):3201 [FREE Full text] [doi: [10.3390/ijerph16173201](#)] [Medline: [31480718](#)]
20. Liu X, Liu X, Sun J, Yu NX, Sun B, Li Q, et al. Proactive suicide prevention online (PSPO): machine identification and crisis management for Chinese social media users with suicidal thoughts and behaviors. *J Med Internet Res* 2019 May 08;21(5):e11705 [FREE Full text] [doi: [10.2196/11705](#)] [Medline: [31344675](#)]

Abbreviations

WHO: World Health Organization

Edited by G Eysenbach; submitted 28.03.21; peer-reviewed by M Aljassas, R Treskes; comments to author 03.05.21; revised version received 11.08.21; accepted 01.10.21; published 11.11.21.

Please cite as:

Chen W, Boggero A, Del Puente G, Olcese M, Prestia D, Jahrami H, Chalghaf N, Guelmami N, Azaiez F, Bragazzi NL

Googling for Suicide-Content and Quality Analysis of Suicide-Related Websites: Thematic Analysis

JMIR Form Res 2021;5(11):e29146

URL: <https://formative.jmir.org/2021/11/e29146>

doi: [10.2196/29146](#)

PMID: [34689118](#)

©Wen Chen, Andrea Boggero, Giovanni Del Puente, Martina Olcese, Davide Prestia, Haitham Jahrami, Nasr Chalghaf, Noomen Guelmami, Fairouz Azaiez, Nicola Luigi Bragazzi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.11.2021. 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

Conducting Health Literacy Research With Hard-to-Reach Regional Culturally and Linguistically Diverse Populations: Evaluation Study of Recruitment and Retention Methods Before and During COVID-19

Genevieve Perrins¹, BSocSci; Tabassum Ferdous¹, MBBS, GDip, MHS, PhD; Dawn Hay¹, BHS, MNA, RN, RM; Bobby Harreveld², DipT&AS, DipTch, BEd, MEd, PhD; Kerry Reid-Searl³, BHLthSc, MCLinEd, PhD, RN, RM

¹Central Queensland Multicultural Association, CQUniversity Rockhampton North, Rockhampton, Australia

²Centre for Research in Equity, Advancement of Teaching & Education, School of Education and the Arts, Central Queensland University, Rockhampton, Australia

³School of Nursing, Midwifery & Social Science, Central Queensland University, Rockhampton, Australia

Corresponding Author:

Genevieve Perrins, BSocSci
Central Queensland Multicultural Association
CQUniversity Rockhampton North
Room 31, Building 41 Buzacott Circle
554-700 Yaamba Road, Norman Gardens
Rockhampton, 4701
Australia
Phone: 61 423853809
Email: evie.perrins@cqma.org.au

Abstract

Background: In health research, culturally and linguistically diverse (CALD) health care consumers are cited as hidden or hard to reach. This paper evaluates the approach used by researchers to attract and retain hard-to-reach CALD research participants for a study investigating health communication barriers between CALD health care users and health care professionals in regional Australia. As the study was taking place during the COVID-19 pandemic, subsequent restrictions emerged. Thus, recruitment and retention methods were adapted. This evaluation considered the effectiveness of recruitment and retention used throughout the pre-COVID and during-COVID periods.

Objective: This evaluation sought to determine the effectiveness of recruitment and retention efforts of researchers during a study that targeted regional hard-to-reach CALD participants.

Methods: Recruitment and retention methods were categorized into the following 5 phases: recruitment, preintervention data collection, intervention, postintervention data collection, and interviews. To compare the methods used by researchers, recruitment and retention rates were divided into pre-COVID and during-COVID periods. Thereafter, in-depth reflections of the methods employed within this study were made.

Results: This paper provides results relating to participant recruitment and retainment over the course of 5 research phases that occurred before and during COVID. During the pre-COVID recruitment phase, 22 participants were recruited. Of these participants, 15 (68%) transitioned to the next phase and completed the initial data collection phase. By contrast, 18 participants completed the during-COVID recruitment phase, with 13 (72%) continuing to the next phase. The success rate of the intervention phase in the pre-COVID period was 93% (14/15), compared with 84.6% (11/13) in the during-COVID period. Lastly, 93% (13/14) of participants completed the postintervention data collection in the pre-COVID period, compared with 91% (10/11) in the during-COVID period. In total, 40 participants took part in the initial data collection phase, with 23 (58%) completing the 5 research phases. Owing to the small sample size, it was not determined if there was any statistical significance between the groups (pre- and during-COVID periods).

Conclusions: The success of this program in recruiting and maintaining regional hard-to-reach CALD populations was preserved over the pre- and during-COVID periods. The pandemic required researchers to adjust study methods, thereby inadvertently contributing to the recruitment and retention success of the project. The maintenance of participants during this period was due

to flexibility offered by researchers through adaptive methods, such as the use of cultural gatekeepers, increased visibility of CALD researchers, and use of digital platforms. The major findings of this evaluation are 2-fold. First, increased diversity in the research sample required a high level of flexibility from researchers, meaning that such projects may be more resource intensive. Second, community organizations presented a valuable opportunity to connect with potential hard-to-reach research participants.

(*JMIR Form Res* 2021;5(11):e26136) doi:[10.2196/26136](https://doi.org/10.2196/26136)

KEYWORDS

health literacy; cultural and linguistic diversity; COVID-19; health care barriers; hard-to-reach research participants; regional Australia; health literacy profiles; literacy

Introduction

Background

Low consumer health literacy levels represent an ongoing challenge to health services. Lower health literacy impacts an individual's ability to access appropriate health services, comprehend medical instruction, and manage their own health [1]. In this way, individuals with low health literacy may forego preventative and proactive management of their health, leading to increased hospitalizations and presentation at emergency departments [1]. Despite the systemic support for promoting active and engaged health care consumers [2], the reality is that certain population subgroups have more difficulty in communicating their needs, engaging with health care professionals, and navigating the health care system. Culturally and linguistically diverse (CALD) health care consumers are at greater risk of lower health literacy levels due their social, cultural, geographic, and economic contexts [3].

Despite the importance of researching the experiences of CALD health care users, this demographic is often cited as hidden or hard to reach [4]. The reasons for this categorization are multiple; migrants may be socially disconnected, may be vulnerable, may have a fear of discrimination, may misinterpret or misunderstand the research process, or may generally distrust researchers [4]. Therefore, there are significant difficulties in recruiting and retaining such a research cohort. In the pursuit of meaningful research in this area, this paper provides valuable insights into the recruitment and retainment of traditionally hard-to-reach research populations [4].

During the research project, COVID-19 emerged. As a result, recruitment and retention methods had to be adapted to address corresponding lockdowns and restrictions. Researchers analyzed the recruitment and retention methods by breaking down the participation process into 5 distinct phases and comparing pre-COVID and during-COVID approaches for each phase.

These phases include recruitment, preintervention data collection, intervention, postintervention data collection, and interviews.

Objectives

The objective of this study was to provide insights into recruitment and retention methods used during a study targeting a highly diverse hard-to-reach CALD sample in a regional city in Australia.

Methods

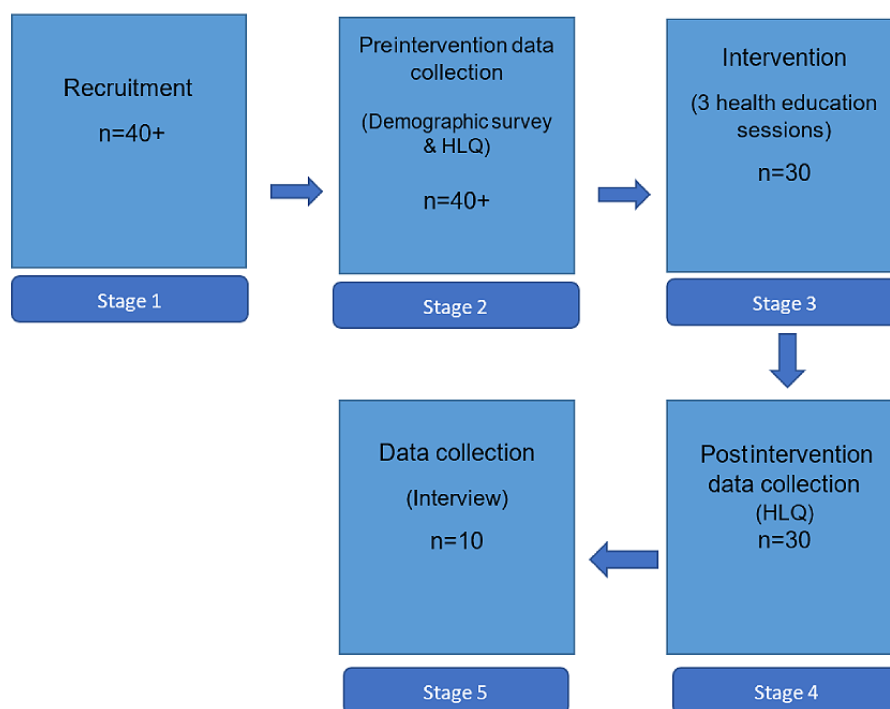
Study Design

The research project was conducted by a community organization based in the central Queensland city of Rockhampton. The organization, Central Queensland Multicultural Association, offers an eclectic range of community programs aimed at overcoming social isolation and promoting integration and engagement. This pilot project was funded by the Australian Government through the Department of Health, for the purposes of (1) establishing an evidence-based health literacy profile for CALD populations residing within the Rockhampton area and (2) examining the efficacy of community-based education sessions as a health intervention option to address ongoing health literacy issues identified from health literacy profiling outcomes.

Participation

While the study sought input from both health care consumers and professionals to establish a health literacy profile, this paper specifically reflects on the difficulties and successes of recruiting CALD health care consumers (herein referred to as participants).

As illustrated in [Figure 1](#), the process of consumer participation included a recruitment phase, preintervention data collection, an intervention phase, and postintervention data collection. Thereafter, 10 participants were invited to give interviews to provide further qualitative data.

Figure 1. Target number of research participants at different phases of the project. HLQ: Health Literacy Questionnaire.

So as not to place heavy restrictions on recruitment, consumer participant criteria were left sufficiently broad. Participants were required to be over the age of 18 years, have moved to Australia from another country, be Medicare eligible, and have adequate English ability.

Quantitative Research Tools

The research program utilized the prevalidated Health Literacy Questionnaire (HLQ) designed by Osborne et al [5]. The HLQ aims to test consumers' self-reported capability for conducting health literacy tasks. The HLQ includes 21 questions that feature Likert-scale responses.

Qualitative Research Tools

Researchers designed an interview schedule that reflected the domains of health literacy set out within the quantitative portion of the study. The schedule included 11 questions relating to the HLQ domains, which were designed as a start point for conducting semistructured interviews.

Intervention

The program's intervention involved 3 education sessions delivered by a researcher. The intervention was designed to (1) inform CALD health consumers about their role in interactions with health care providers, (2) increase consumer understanding and manage expectations of clinical appointments, and (3) inform participants about local health system navigation. Within the education sessions, additional resources were utilized to aid participant understanding, including PowerPoint, as well as videos and resource sheets published by government agencies.

Evaluation Tools

Participants were observed in their dealings with researchers. Researchers made note of participant behavior, including

avoidance, requests to withdraw, and connection-seeking, as well as preferences associated with location and group makeup. These interactions provided important insights into participants' feelings toward the research process, including (1) the perceived significance of the research to the individual and the community, (2) whether it was considered a justified use of time, (3) the need for having social support during the research process, and (4) distrust or anxiety.

These observations were recorded in a research journal and reflected on in meetings between researchers, as well as in discussions with community leaders. The use of observations was a necessary data collection method as researchers had to adjust recruitment and retention methods in response to these observations.

In addition to observations, the program sought participant feedback in order to evaluate the effectiveness of the program's recruitment and retention methods. Data collection methods included the use of an online survey or interview questions. The use of either of these data collection methods was determined by the participants' levels of involvement in the project. For example, if a participant had agreed to be interviewed as part of the research project, researchers would include evaluative questions as part of the interview. However, if participants had not taken part in an interview, had initially signed up but later withdrawn, or failed to complete all the steps for participation, they were sent an online survey. Both data collection methods queried the same potential issues associated with recruitment or retention. These included (1) logistical issues such as transport and childcare, (2) time constraints, (3) the impact of session size and group makeup, and (4) the participation processes.

Evaluation Analysis

Researchers analyzed the responses provided by participants in their interviews and online surveys in order to gather information on what barriers were faced by participants during the research process.

In addition to direct feedback and observation, researchers analyzed the dropout rates of participants within the 5 research phases. The recruitment phase refers to “the dialogue which takes place between an investigator and a potential participant prior to the initiation of the consent process” [6]. The process includes the identification of potential participants, filtration, and discussion of participation steps. The preintervention data collection phase refers to participants completing necessary paperwork to take part in the study. This paperwork included a program consent form, a demographic survey, and a preintervention HLQ. The intervention stage refers to the 3 education sessions (S1, S2, and S3). The postintervention data collection phase refers to the collection of completed HLQs following the 3 education sessions. The interview phase refers

to the completion of a one-on-one semistructured interview with researchers.

As each of these participation phases occurred within pre-COVID and during-COVID periods, researchers were able to adapt recruitment and retention methods, therefore providing a basis for a comparative analysis. Researchers then compared the dropout rates for each participation phase within each designated period (pre-COVID or during-COVID period).

Ethics

Engagement with health care consumers was approved on an ethical basis by the Central Queensland Hospital and Health Service Human Research Ethics Committee (reference: LNR/2019/QCQ/57544).

Results

Table 1 provides a breakdown of the research project’s recruitment and retention results according to the COVID period and research phase.

Table 1. Number of participants before and during COVID-19 restrictions at different phases of the study.

Period	Recruitment (N=40), n (rate)	Initial data collection: demographic and HLQ ^a (N=28), n (%)	Intervention		Postintervention data collection: HLQ (N=23), n (%)	Interview (N=10), n
			Sessions 1 and 2 (N=28), n	Session 3 (N=25), n (%)		
Before COVID-19 restriction (November 30, 2019, to January 30, 2020)	22 (2.9/week)	15 (68%)	15	14 (93%)	13 (93%)	0
During COVID-19 restriction (January 31, 2020, to August 7, 2020)	18 (0.66/week)	13 (72%)	13	11 (85%)	10 (91%)	10

^aHLQ: Health Literacy Questionnaire.

Discussion

Recruitment

A major recruitment success was the diversity of the research sample. The 40 participants who took part had 15 separate

countries of origin. Participant country of origin and world region are presented in Table 2.

Table 2. Participant profile by world region and country of origin.

World region	Country	Participants (N=40), n (%)
Southern Asia	Nepal	8 (20)
Southern Asia	Bangladesh	6 (15)
Southern Asia	Sri Lanka	5 (13)
Middle East	Afghanistan	4 (10)
Mainland South-East Asia	Myanmar	4 (10)
Melanesia	Papua New Guinea	4 (10)
South America	Brazil	1 (3)
Chinese Asia (includes Mongolia)	China	1 (3)
Northern Europe	Denmark	1 (3)
Polynesia (excludes Hawaii)	Fiji	1 (3)
Western Europe	Germany	1 (3)
Southern and East Africa	South Africa	1 (3)
Melanesia	Solomon Islands	1 (3)
Mainland South-East Asia	Thailand	1 (3)
Mainland South-East Asia	Vietnam	1 (3)

The greatest concentration of Queensland's CALD health care users is in Brisbane; as of 2018, 30.6% of the capital's population had a CALD background. However, there are increasing levels of diversity among the regions [7]. Rockhampton's CALD population has been growing steadily; in 2011, 8.7% of the regional city's population was CALD, compared with 9.4% in 2016 [7]. Further, as of 2016, 6.8% of Rockhampton's population spoke a language other than English at home [8]. However, this is not to say that Rockhampton's CALD population is necessarily homogenizing. Owing to governmental policy, which promotes migration to the region [9], the city's migrant population is likely to become increasingly diverse. In contrast to studies of CALD health care consumers in metropolitan areas, research in regional areas must account for a significant amount of cultural diversity within the sample. Therefore, recruitment and retention methods that can be adapted to multiple cultural backgrounds may be more successful. In the case of this research program, the ability to recruit and retain such a diverse sample represents a success.

Pre-COVID Period

The program used purposive, convenience, and snowballing methods to recruit participants. Beyond participants who were recruited via broader means (eg, public community days discussed further under the Intervention subheading), researchers utilized immediate networks to gauge the interest of potential participants. The use of established community networks was of great benefit to researchers; it required a lesser degree of time and resources to locate potential participants, and researchers had already established a level of rapport through other dealings with community members.

While snowballing and convenience sampling methods in this project were largely effective in recruiting participants, overreliance on established networks can be problematic for the research objectives. Overreliance on such a network risks

introducing self-selection bias, as willing participants may be well settled, educated, nonisolated, and interested in research [4]. In spite of this limitation, ethnic minorities are often difficult to recruit using traditional sampling methods owing to suspicion of research/researchers, vulnerability, social isolation, and stigma [4]. Although an acknowledged bias, other options for reaching hidden populations is yet to be discovered, and as such, limitations need to be endured in order to pursue much needed research in this area. Within this study, care was taken to ensure that snowballing starting points were as diversified as possible. As illustrated in Table 2, initial recruitment and potential networks were selected by country of origin to ensure that broad geographic regions were accounted for.

To further diversify participant background, researchers also attempted to broaden the recruitment method by engaging with other community institutions and their leaders. Such organizations included language groups, and cultural and social organizations, as well as places of worship. There was significant potential in connecting with such organizations, and if successful, there was access to a high number of diverse participants. However, researchers found this approach to be ineffectual. While organizations had expressed an interest in connecting contacts with researchers, in each case, these attempts were unsuccessful. This approach was therefore found to be highly bureaucratic and resource intensive.

During-COVID Period

With COVID restrictions ceasing many day-to-day activities that researchers had previously used as a gateway to recruitment, as well as the difficulty in recruiting participants via institutional leaders, researchers sought out community leaders instead as the instigation point for recruitment. In this paper, institutional leaders refer to those who are paid, who hold positions of authority designated to them by way of a constitution/contract, or who hold sway over large groups of people not individually

known to them. It is a formal position. Community leaders, by contrast, are informal positions not held within an institution, but they instead are figures who hold authority by way of their connection with individuals residing in the community. Often, community leaders are associated with specific religious, ethnic, or cultural groups.

It was recognized that a direct approach to individuals within the CALD community was not effective. It resulted in participant reluctance due to fear of the research process and researchers, as well as stigma. Instead, researchers decided to approach cultural and religious groups. In doing so, researchers would reach out to community leaders to gauge interest in the project. The project used the established network of the organization to identify and recruit community leaders. For example, the diversity of the organization's management committee and workforce allowed researchers to tap into hidden South American populations, recruit socially isolated participants from within the local Mosque, and access a women's group based around cultural diversity and inclusion activities. Further, the organization itself represented a drawcard to ethnic minorities, as these groups would seek out assistance relating to the organization's broader function. Their attendance at the organization allowed researchers to approach for discussing the research program and participation.

Community leaders acted as research gatekeepers; they would explain the process of participation to potential participants, provide emotional support, and act as mediators between the researchers and participants. The use of gatekeepers in research is largely successful, particularly where the subject is deemed significant within the community [10]. The findings of this program were that the use of cultural/religious gatekeepers is highly effective.

In addition to the use of cultural gatekeepers, approaching participants within their cultural groups overcame recruitment issues. Such an approach worked well during COVID-19 restrictions, when opportunities for researchers to themselves establish contact and rapport were diminished. Researchers felt that community inclusion overcame issues of participant self-assuredness and made participation seem less daunting. Participants were more forthcoming in speaking of their experiences, they had a chance to build rapport with researchers in a less threatening environment, and it provided a level of anonymity for participants. From the researchers' perspective, the increase in willingness from the side of the participants was attributable to both the use of intimate cultural groups and the use of cultural gatekeepers.

The shift from recruiting participants face to face at events to a purely network-based recruitment method meant that enlistment became more targeted. Difficulties in recruiting CALD participants were seemingly compounded when potential participants encountered an Anglo-Australian researcher. The researchers tasked with data collection were 2 women (one had an Anglo-Australian background, and other had a Bangladeshi background). Observational evidence acquired by the researchers suggested that the meeting of researcher and participant backgrounds was significant. In fact, in cases where age or gender disparities between an Anglo-Australian researcher and

potential participants were minimal, the participants still gravitated toward researchers twice their age, but with similar religious or cultural backgrounds. The success rate of recruitment was variable, with more success when participants were faced with a researcher from a migrant background. This finding corresponds with previous literature, which identifies perceived power differentials as a major factor in the dearth of CALD individual participation in health research [11-13]. In response to these difficulties and the issues outlined in the research, adjustments to recruitment were made to ensure that the CALD researcher represented the "face" of the program, with participants enjoying maximum contact with this researcher where possible.

Preintervention Data Collection

Pre-COVID Period

In order to minimize the dropout rate between recruitment and data collection, those who agreed to participate during the recruitment phase were presented immediately with data collection materials. This approach worked well when the research organization had agreements with cultural groups for the use of facilities. For example, the research organization would allow the use of rooms for language classes. Often when these classes were held, members of cultural groups would congregate. Researchers used the opportunity to recruit these participants and present them immediately with data collection materials. For establishing a health literacy profile for this region, this method was highly successful; 14 participants were initially approached, with 8 being eligible and successfully completing the preintervention data collection phase. Participants were happy to be approached and document their experience; however, this did not necessarily translate to high levels of participant retention. In the pre-COVID period, 8 participants were recruited in this fashion; however, only 3 of those were retained and eventually took part in the intervention phase. The dropout rate in this instance was hypothesized as participants being reluctant to return to complete the intervention. Researchers felt that for this research cohort, making the research process as seamless as possible may be more attractive than requiring multiple distinct steps to complete participation.

During-COVID Period

As normal activities conducted by the organization ceased under COVID restrictions, routine data collection processes were amended to ensure the feasibility of this research phase. The researchers' recruitment pitch had to highlight the importance of the research objectives to not only the individuals, but also their broader community [14]. In the pre-COVID data collection phase, while this was still an important factor to be included, participants were easily swayed by the limited amount of time and resources necessary to participate. Throughout the during-COVID data collection phase, researchers were required to follow-up with participants to enquire about any difficulties, to arrange collection of the paperwork, and to discuss the next stages.

Even with the increased barriers to participation that COVID-19 posed, a continuance rate of 72% (13/18) during this period

indicates that researchers successfully adapted their data collection approach so that it suited CALD research participants. This was done by offering a high level of flexibility, such as the collection of paperwork from residences or places of work (at the participant's request), the use of digital platforms to deliver and receive completed questionnaires, and the ability of researchers to gather data collection tools at the upcoming intervention phase. Despite the level of flexibility, some participants were difficult to retain, and ceased contact with researchers. These participants did not provide feedback to researchers; however, it is believed that participant withdrawal in this period was due to potential time constraints and unawareness of the research process. Unfortunately, in cases where participants had ceased contact, the requirements of participation could not be explained.

Intervention

Pre-COVID Period

The original approach was to hold sizeable community sessions, whereby a large number of participants could be provided with the first 2 sessions (S1 and S2) of the intervention. This approach not only satisfied program funding arrangements with respect to community engagement, but also allowed for minimal resource requirements in terms of staffing and time. It was initially anticipated that the program would require 3 to 4 of these large gatherings to collect consumer participant numbers. The first intervention took place on the community health day. The health day was advertised to the broader community, and no restrictions were placed on attendance. Researchers had arranged for local health professionals to attend and provide free health checks. Resource packs were provided to attendees with local health information, as well as navigational aids for the region's health service. Further, the research organization arranged for childcare services for the duration of the health day. Researchers took the opportunity to identify potential participants who attended. Those who agreed to take part were presented with data collection materials before the intervention took place. Each of these sessions was held onsite, at the organization's office.

Embedding the intervention within the community health day represented an effective way to engage potential participants. Many community members attended the intervention as part of the community health day but did not necessarily agree or were eligible to participate in the research project. The total number of attendees at community days was 23, with 10 being eligible and agreeing to participate. Of the 13 who did not participate, 7 were ineligible. Therefore, 6 were eligible but did not agree to participate. Researchers hypothesized that a lack of awareness of the research process and aims was to blame for the reluctance of eligible nonparticipants. However, because the intervention was delivered to sizable groups, it was not possible for researchers to identify potential problems and reassure reluctant nonparticipants. Of the 10 participants who agreed to participate during the community health days, 9 went on to complete the intervention in its entirety. Despite this reasonable success, subsequent attempts to organize the third portion of the intervention (S3) in this way were met with participant

reluctance, and it was not possible to replicate the earlier success.

Researchers considered that this reluctance was due to education sessions that were presented too formally and with large culturally diverse groups. The formulation of education sessions was done with oversight from health science academics and staff from the health service. However, due to restraints associated with a pilot study (discussed further in the Limitations subsection), the education sessions were not formulated with community feedback. Subsequently, as the project responded to COVID restrictions and the delivery of sessions was to smaller culturally homogenous groups, feedback from participants suggested that informal delivery, including participant discussion and question time, was more impactful and more aligned with participant expectations.

The pre-COVID intervention was always delivered face-to-face. The choice to deliver the program directly to participants is symptomatic of a broader organizational approach to overcoming social isolation and enhancing integration. The success of the face-to-face approach was highly dependent on the individual. For example, some participants sought out the research program as a way to establish contact with the organization and community in order to overcome social isolation. These participants gained more from direct face-to-face contact with researchers than they would have from an online program. However, in other cases where participants were sufficiently integrated, had established themselves within the broader community, and had higher employment or familial commitments, face-to-face interaction represented a barrier to participation. Researchers were aware of the individual circumstances of each participant, and as such, would regularly hold sessions outside business hours, such as on weeknights or weekends, and they employed childcare services to ensure women could attend and were mentally present during the intervention. Both approaches were effective in overcoming barriers. Due to the effectiveness of these approaches, they were again utilized by the research team in the during-COVID intervention phase.

During-COVID Period

The emergence of COVID-19 brought with it restrictions on face-to-face events, which required alterations to session delivery within the program. Researchers initially minimized the number of participants present at each intervention session and shifted the venue from the research organization to one of participant choosing. The session size was 10 or over pre-COVID but had to be restricted to no more than five in each session. Researchers found that participant willingness to partake in the entire intervention process increased during this period. Participants indicated that smaller and more intimate sessions gave them the opportunity to speak about their experiences and allowed researchers to refine the intervention to suit particular individuals and their needs. While the numbers during this period may suggest a decrease in participation, subsequent discussions with participants suggested that sessions taking place under this method were more meaningful to each participant. Further, in holding sessions in spaces of participant choosing, the program effectively overcame potential barriers

experienced by participants such as work commitments, transport, childcare, and time constraints.

During the height of restrictions, in which contact with persons outside the household was disallowed, the intervention continued via online platforms. Sessions 1 and 2 were delivered by a researcher via Zoom (Zoom Video Communications). The third session was delivered either by Zoom or via a prerecorded session uploaded to YouTube (Google Inc). This arrangement allowed for researchers to connect with participants, thereby increasing participant opportunity to speak of their experiences and ask questions while adhering to public health restrictions. In some cases, participants benefitted greatly from the flexibility. However, while the project maintained targeted numbers during this period, the relationship fostered online, compared with face-to-face, is more superficial. Research participants of diverse backgrounds have a greater need for social comfort in a research context [11]. Therefore, such needs risk not being met when researchers make participation a solely online activity.

Postintervention Data Collection

Pre-COVID Period

Following successful completion of the intervention, participants were required to fill in a postintervention HLQ. Initially, researchers posted the HLQ to participants and requested them to return the paperwork either to the organization's office, via post, or email. Researchers attempted to give sufficient consideration for individual circumstances; however, expecting some participants to return the paperwork increased barriers to participation. In particular, researchers noticed significant participant fatigue by this stage of the program, therefore requiring researchers to provide further flexibility to maintain participant numbers. To provide this flexibility, researchers collected the data collection materials from a location and at a time requested by the participants. The finding of the research team was that the initial approach in requiring participants to return their paperwork was largely ineffective. This is because it inadvertently placed further restrictions on participation, such as the need for transport and childcare, as well as needing to fit this responsibility between employment or familial obligations.

During-COVID Period

The collection of hard-copy postintervention data during COVID restrictions, particularly for participants who required assistance, was laborious. However, in cases where participants had been identified as sufficiently literate (both digitally and with forms) researchers used Google Forms as a way to disperse questionnaires. It was considered that reworking the paperwork into a Google Forms format was beneficial in that it may have appeared less daunting than the layout of the tangible form, while minimizing COVID-related and traditional barriers to participation. When virus restrictions allowed, the research team resumed collection of tangible forms.

The collection of tangible forms was preferable as it gave an opportunity to assist participants who faced difficulties in filling out the questionnaire. The finding of this project was that the presentation and language of the HLQ involved some difficulties. First, some participants suggested that they had not encountered a survey of this type (Likert scale) and mentioned

that they were unsure about how to correctly fill in the form. Second, despite participants speaking English, some encountered difficulties with the language used in the data collection material.

Interviews

After completing the quantitative aspect of the study, 10 participants were invited from the sample to take part in semistructured interviews. There were no further selection criteria to take part in the interview phase, instead researchers were opportunistic. The initial study design called for all quantitative data to be collected before interviews could take place. This was so that participants could build upon the answers provided in the survey, as well as supply the project with feedback about their experiences with the pilot program. For these reasons, the interview phase of the project only occurred throughout the during-COVID period, and therefore, success rates cannot be compared.

The researchers approached interviews on a flexible basis, often offering participants the choice of an online Zoom interview or an interview at a location of their choosing. Time frames were also flexible, with researchers making appointments according to participants' availability (eg, weekends and nights). As the study only required 10 participants for the interviews, the slots were filled quickly and easily.

Evaluation Limitations

Community-based interventions designed to address health literacy within population subgroups may undergo a process of needs assessment before intervention co-design and implementation [15]. Within this pilot study, it was not feasible to establish an intervention off the back of rigorous community engagement, instead this program sought input from the health service and academics within health sciences. This is a noted limitation; it is preferable to incorporate input from the target population in order to ensure intervention methods are appropriate and tailored [15]. If education sessions were formulated with more feedback from the intended population, participant recruitment and retention might have been enhanced.

Further, this pilot project did not employ translators or interpreters for participants. Researchers considered that the use of accredited interpreters and translators would diversify the sample by removing the need for participants to be proficient in English, as well as ensuring that interpretations were true to their original meaning. Considering the diversity of the research sample, the cost of translating materials and employing interpreters within this pilot study was prohibitive. To mitigate difficulties, researchers provided assistance and clarification when requested to do so. Researcher involvement with completion of the questionnaire has the potential to introduce bias; however, such a limitation had to be endured to be sure that certain participants understood what was asked and how to properly record their answers.

Conclusions

This paper documented the efforts of researchers to recruit and retain CALD participants before and during COVID-19 restrictions. While it was anticipated by researchers that COVID

would negatively impact the research program, alternative action in response to COVID actually assisted researchers in identifying the strengths and weaknesses of prior methods and gave an opportunity for researchers to evaluate the efficacy of different methods. The unforeseen pandemic inadvertently forced researchers to rethink the initial study design, which benefited the research aims.

There are a number of takeaways from this program. First, the sample used by this project was highly diverse. The sample included 40 individuals having 15 countries of origin. While other studies have targeted and discussed recruitment difficulties and successes as they relate to specific ethnic groups, there is a shortage of information relating to the recruitment of a research sample that involves a high level of cultural diversity. This study identified the way in which the term CALD obscures the true diversity of regional-based migrant populations and provided necessary insights into the way in which such a group can be incorporated into health research. Lastly, this project highlighted the way in which community-based organizations have a discernible role in connecting hard-to-reach participants with research projects.

In a study such as this, difficulties in recruitment and retention may have spurred similar actions irrespective of the occurrence of the pandemic. Therefore, in similar studies conducted outside the impact of the pandemic, recruitment and retention may look similar, that is, first attempts may aim to exhaust traditional methods, such as large-scale recruitment via community health days, large intervention sessions, and use of institutions to recruit community members, before turning toward tailored culture-specific bottom-up approaches. The initial success of the pre-COVID period, followed by recruitment stagnation, as well as the ongoing success of during-COVID procedures, suggests that a high level of diversity in the research sample requires a high level of diversity in recruitment and retention methods. Therefore, the keys to ensuring successful recruitment and retention are to provide flexibility; respond to individual and group circumstances; reflect on specific environmental, geographic, and socioeconomic barriers; and mitigate issues. Considering the heightened requirement for flexibility, a major takeaway from this study is that research focusing on CALD samples is inherently more time and resource intensive. Therefore, future studies that target such a population will need to factor this into the study design and cost.

Acknowledgments

We recognize the valuable contributions of the project's partners. These include Central Queensland University that provided guidance, and Central Queensland Hospital and Health Service that provided strategic support. This pilot project was funded by the Australian Commonwealth Government, through the Department of Health.

Conflicts of Interest

None declared.

References

1. Australia's Health 2018. Australian Institute of Health and Welfare. 2018. URL: <https://www.aihw.gov.au/reports/australias-health/australias-health-2018/contents/table-of-contents> [accessed 2021-10-16]
2. Destination 2030: Great Care for Central Queenslanders. Central Queensland Hospital Health Service. URL: https://www.health.qld.gov.au/_data/assets/pdf_file/0037/669772/destination-2030-full.pdf [accessed 2021-10-16]
3. Kreps GL, Sparks L. Meeting the health literacy needs of immigrant populations. *Patient Educ Couns* 2008 Jun;71(3):328-332. [doi: [10.1016/j.pec.2008.03.001](https://doi.org/10.1016/j.pec.2008.03.001)] [Medline: [18387773](https://pubmed.ncbi.nlm.nih.gov/18387773/)]
4. Ellard-Gray A, Jeffrey NK, Choubak M, Crann SE. Finding the Hidden Participant. *International Journal of Qualitative Methods* 2015 Dec 17;14(5):1609406915621420. [doi: [10.1177/1609406915621420](https://doi.org/10.1177/1609406915621420)]
5. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). *BMC Public Health* 2013 Jul 16;13:658 [FREE Full text] [doi: [10.1186/1471-2458-13-658](https://doi.org/10.1186/1471-2458-13-658)] [Medline: [23855504](https://pubmed.ncbi.nlm.nih.gov/23855504/)]
6. Patel MX, Doku V, Tennakoon L. Challenges in recruitment of research participants. *Adv. Psychiatr. Treat* 2018 Jan 02;9(3):229-238 [FREE Full text] [doi: [10.1192/apt.9.3.229](https://doi.org/10.1192/apt.9.3.229)]
7. Diversity Figures June 2018. The State of Queensland. 2018. URL: <https://www.cyjma.qld.gov.au/resources/dcsyw/multicultural-affairs/multicultural-communities/diversity-figures-report-2018.pdf> [accessed 2021-10-16]
8. 2016 Census QuickStats. Australian Bureau of Statistics. URL: https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/LGA36370?opendocument [accessed 2020-09-13]
9. Garad R, Waycott L. The Role of Health Literacy in Reducing Health Disparities in Rural CaLD Communities. The National Rural Health Alliance. URL: http://www.ruralhealth.org.au/13nrhc/images/paper_Garad%2C%20Rhonda_Waycott%2C%20Lauren.pdf [accessed 2020-07-13]
10. Renert H, Russell-Mayhew S, Arthur N. Recruiting Ethnically Diverse Participants into Qualitative Health Research: Lessons Learned. *TQR* 2015 Jan 14;18(12):1-13 [FREE Full text] [doi: [10.46743/2160-3715/2013.1542](https://doi.org/10.46743/2160-3715/2013.1542)]

11. Ahmed A, Vandrevalla T, Hendy J, Kelly C, Ala A. An examination of how to engage migrants in the research process: building trust through an 'insider' perspective. *Ethn Health* 2019 Nov 13:1-20. [doi: [10.1080/13557858.2019.1685651](https://doi.org/10.1080/13557858.2019.1685651)] [Medline: [31722534](https://pubmed.ncbi.nlm.nih.gov/31722534/)]
12. Hughson J, Woodward-Kron R, Parker A, Hajek J, Bresin A, Knoch U, et al. A review of approaches to improve participation of culturally and linguistically diverse populations in clinical trials. *Trials* 2016 May 26;17(1):263 [FREE Full text] [doi: [10.1186/s13063-016-1384-3](https://doi.org/10.1186/s13063-016-1384-3)] [Medline: [27229153](https://pubmed.ncbi.nlm.nih.gov/27229153/)]
13. Fête M, Aho J, Benoit M, Cloos P, Ridde V. Barriers and recruitment strategies for precarious status migrants in Montreal, Canada. *BMC Med Res Methodol* 2019 Feb 26;19(1):41 [FREE Full text] [doi: [10.1186/s12874-019-0683-2](https://doi.org/10.1186/s12874-019-0683-2)] [Medline: [30808301](https://pubmed.ncbi.nlm.nih.gov/30808301/)]
14. Ferris EB, Wyka K, Evenson KR, Dorn JM, Thorpe L, Catellier D, et al. Recruitment and Retention Strategies for Community-Based Longitudinal Studies in Diverse Urban Neighborhoods. *JMIR Form Res* 2021 Mar 24;5(3):e18591 [FREE Full text] [doi: [10.2196/18591](https://doi.org/10.2196/18591)] [Medline: [33759799](https://pubmed.ncbi.nlm.nih.gov/33759799/)]
15. Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga L, et al. Co-design of mHealth Delivered Interventions: A Systematic Review to Assess Key Methods and Processes. *Curr Nutr Rep* 2016 Jul 4;5(3):160-167. [doi: [10.1007/s13668-016-0165-7](https://doi.org/10.1007/s13668-016-0165-7)]

Abbreviations

CALD: culturally and linguistically diverse

HLQ: Health Literacy Questionnaire

Edited by G Eysenbach; submitted 06.12.20; peer-reviewed by A Parker, H Imeri; comments to author 09.02.21; revised version received 29.05.21; accepted 20.09.21; published 02.11.21.

Please cite as:

Perrins G, Ferdous T, Hay D, Harreveld B, Reid-Searl K

Conducting Health Literacy Research With Hard-to-Reach Regional Culturally and Linguistically Diverse Populations: Evaluation Study of Recruitment and Retention Methods Before and During COVID-19

JMIR Form Res 2021;5(11):e26136

URL: <https://formative.jmir.org/2021/11/e26136>

doi: [10.2196/26136](https://doi.org/10.2196/26136)

PMID: [34581673](https://pubmed.ncbi.nlm.nih.gov/34581673/)

©Genevieve Perrins, Tabassum Ferdous, Dawn Hay, Bobby Harreveld, Kerry Reid-Searl. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 02.11.2021. 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

A Case Study of an SMS Text Message Community Panel Survey and Its Potential for Use During the COVID-19 Pandemic

Lilian Chan¹, MBBS, MPH; Nouhad El-Haddad², BSc, MND, PhD; Becky Freeman¹, BSc, MSc, PhD; Blythe J O'Hara¹, BA, BSW, MAppSci, PhD; Lisa Woodland³, BSc (Hons); Ben Harris-Roxas⁴, BSW, MPASR, PhD

¹Prevention Research Collaboration, Sydney School of Public Health and Charles Perkins Centre, The University of Sydney, Camperdown, Australia

²Centre for Primary Health Care and Equity, Faculty of Medicine, University of New South Wales, Sydney, Australia

³Population and Community Health, South Eastern Sydney Local Health District, Darlinghurst, Australia

⁴School of Population Health, Faculty of Medicine, University of New South Wales, Sydney, Australia

Corresponding Author:

Lilian Chan, MBBS, MPH

Prevention Research Collaboration

Sydney School of Public Health and Charles Perkins Centre

The University of Sydney

John Hopkins Drive

Camperdown, 2006

Australia

Phone: 61 286277554

Email: lilian.chan@sydney.edu.au

Abstract

During the COVID-19 pandemic many traditional methods of data collection, such as intercept surveys or focus groups, are not feasible. This paper proposes that establishing community panels through SMS text messages may be a useful method during the pandemic, by describing a case study of how an innovative SMS text message community panel was used for the “Shisha No Thanks” project to collect data from young adults of Arabic-speaking background about their attitudes on the harms of waterpipe smoking. Participants were asked to complete an initial recruitment survey, and then subsequently sent 1 survey question per week. The study recruited 133 participants to the SMS text message community panel and the mean response rate for each question was 73.0% (97.1/133) (range 76/133 [57.1%] to 112/133 [84.2%]). The SMS text message community panel approach is not suited for all populations, nor for all types of inquiry, particularly due to limitations of the type of responses that it allows and the required access to mobile devices. However, it is a rapid method for data collection, and therefore during the COVID-19 pandemic, it can provide service providers and policymakers with timely information to inform public health responses. In addition, this method negates the need for in-person interactions and allows for longitudinal data collection. It may be useful in supplementing other community needs assessment activities, and may be particularly relevant for people who are considered to be more difficult to reach, particularly young people, culturally and linguistically diverse communities, and other groups that might otherwise be missed by traditional methods.

(*JMIR Form Res* 2021;5(11):e28929) doi:[10.2196/28929](https://doi.org/10.2196/28929)

KEYWORDS

data collection; mobile phone; short message service; tobacco; COVID-19; survey

Introduction

There is a high level of interest in communities' experiences and needs during the COVID-19 pandemic [1-4], and as people's lived experiences have been heterogeneous, more information is needed to understand what different subgroups within populations have been through. This is especially true of minority populations who are underrepresented in mainstream conversations. Traditional methods of collecting participant

data, such as intercept surveys or focus groups, are not feasible during a pandemic due to physical distancing requirements.

Establishing community panels that provide data through SMS text messaging is a potential method that could be used during the COVID-19 pandemic to provide information for support services, policy planning, and research studies. SMS text message community panels allow longitudinal data collection and involve the recruitment of a sample of the community to form the community panel, and then sending a small number

of survey questions via SMS text messages to the panel participants at regular short time intervals. Panel participants respond to the survey questions by sending a short response back via an SMS text message. This method allows tailoring of language to different community groups, and is particularly suitable to younger people who have been socialized to communicate using this channel. SMS text message surveys have been trialed in health research studies for data collection and have been found to be user-friendly and produce reasonable response rates [5-9].

While there are other noncontact methods of data collection that are useful during the COVID-19 pandemic, such as online surveys [10-13] and online focus groups [14], we propose that SMS text message community panels could be an additional useful tool. This report highlights an example of an SMS text message community panel, and then discusses how the approach could be used during the pandemic.

The “Shisha No Thanks” SMS Text Message Community Panel

Overview

“Shisha No Thanks” is a co-design project that aims to raise awareness about the harms of waterpipe (shisha) smoking among young adults (aged 18-35 years) of Arabic-speaking background in Sydney, New South Wales, Australia [15,16]. To evaluate the campaign, a community panel using SMS text message-delivered survey questions was established to identify changes in attitudes about the harms of waterpipe smoking. The survey questions were specifically designed for this study to measure awareness of project messages and attitudes toward risks of waterpipe smoking, and questions were adapted from the Cancer Institute NSW Tobacco Tracking Survey [17] and the Syrian Center for Tobacco Studies Narghile-Waterpipe Users Survey [18]. As 94% of young people (18-29-year olds) have smartphones [19], SMS text message-delivered survey questions were considered to be a useful way of engaging the target audience.

Recruitment Process

Recruitment advertisements directed people to an online recruitment survey which was built using Qualtrics software.

Textbox 1. Questions from the Shisha No Thanks SMS text message survey (English language version).

How would you rate smoking shisha compared to cigarettes considering its health effects?
1. Same
2. Less harmful
3. More harmful
4. Don't know
What's the main reason(s) you smoke shisha (in a few words):
Have you recently talked to someone (eg, family or friend) about the harms of smoking shisha?
A. Yes
B. No
C. Don't know

The recruitment survey had information about the study, and then asked for demographic information, a mobile phone number, and consent to participate in the study. People who completed the online recruitment survey were then added to the SMS text message community panel database.

To recruit and retain participants, financial reimbursements were provided. The recruitment material explained to participants that they would be compensated for their time with e-gift cards valued at AUD 50 (US \$37) each, which would be sent via SMS text messages at 3 different stages of the study if they answered 75% or more of the questions. To provide context, AUD 50 (US \$37) is equivalent to 6.5% of the minimum weekly wage in this country [20].

SMS Text Message Survey Development

SMS text message community panel members were then sent 1 survey question per week. In total, the study survey consisted of 22 questions—a set of 8 questions that were asked at the beginning of the study period, 6 other questions, and then the initial set of 8 questions were asked again. This was designed for longitudinal follow-up of the cohort, to detect changes in attitudes and awareness about the harms of waterpipe smoking before and after the project.

The survey questions were also set up using Qualtrics, which allows for questions to be sent via SMS text messages. While there are many tools available for SMS text message surveys, Qualtrics was selected for this study, as it is a platform available through research institution licensing, allows for secure data hosting arrangements, and provides a user-friendly process to build SMS text message surveys (as it uses the same interface as the one used to build online surveys). It is worthwhile to note that SMS text message distribution in Qualtrics is an add-on feature, and not part of its standard license.

Each question was set up as an individual survey, and an identification number was assigned to each participant to match his/her responses throughout the study. The survey questions were designed specifically for SMS text messages, with short and concise questions that can be responded to using either multiple choice or short-text answers. Study participants were able to participate in English or Arabic (see [Textboxes 1 and 2](#) for examples of the survey questions in English and Arabic).

Textbox 2. Questions from the Shisha No Thanks SMS text message survey (Arabic language version).



Recruited Community Panel

The study was able to recruit 133 participants for the SMS text message community panel. This was roughly equivalent to the sample size that the project had planned for (n=100 paired responses for each response, anticipating that not all participants would respond to every question).

Community panel participants were recruited through the local community partner's communication channels, including email newsletters and social media pages; through active local community champions who shared the recruitment survey link with their own networks via email, SMS text message, or in person; and through printed flyers with the survey link at community events, such as tertiary education open days. The recruitment survey was available on a tablet device for participants to complete at these events. Local community partners and active community champions were provided with all the recruitment materials that were used to promote the online recruitment survey.

The research team perceived the following factors to be influential in the recruitment and retention of community panel participants to the study: nature of the study by reducing the burden of participation, participants' age range, close engagement with the community during recruitment, participants being financially compensated for their time, and providing

Arabic translations for individuals who do not speak English or prefer to participate in Arabic. Although standard SMS text messaging rates applied to the participants to answer each question, this did not hinder their response rate.

The SMS text message community panel participants' age ranged from 18 to 35 years (mean 25.8 [SD 5.1]), with 64.7% (86/133) being female. In terms of language spoken at home, 12/133 (9.0%) spoke only Arabic, while 87/133 (65.4%) spoke English and Arabic. These demographics were consistent with the target group the research was designed to study. Only 5/133 (3.8%) participants opted to complete the survey in Arabic.

Response Rates

The SMS text message community panel participants received questions on their phone via SMS text messages. To respond to the question, they sent their response by replying to the same number via SMS text message and typing in either a multiple-choice response or a short text. The mean response rate for individual survey questions was 73.0% (97.1/133) (range 76/133 [57.1%] to 112/133 [84.2%]). This response rate is comparable to the rates reported in other studies using SMS text message surveys [6-9]. Table 1 shows the response rates for each question that was asked before and after the project. Response rates for 6 out of 7 questions were lower for the second round.

Table 1. Response rates for each survey question.

Question ^a	Participants who responded (N=133)	
	First round (before project), n (%)	Second round (after project), n (%)
Q1	101 (75.9)	89 (66.9)
Q3	105 (78.9)	87 (65.4)
Q4	103 (77.4)	87 (65.4)
Q5	112 (84.2)	85 (63.9)
Q6	106 (79.7)	93 (69.9)
Q7	76 (57.1)	93 (69.9)
Q8	105 (78.9)	87 (65.4)

^aQ2 has not been included, as it was only sent to participants who answered yes to Q1.

Study Design Challenges

The main challenge encountered was related to the Arabic translation of participant material, including the participant information and consent form, online recruitment survey, and the SMS text message survey questions. An accredited translator translated these materials from English to Arabic. To check for accuracy and content, an Arabic-speaking researcher on the evaluation team compared the translated version with the original English version. However, during the initial recruitment phase, some Arabic-speaking participants who chose to complete the online recruitment survey in Arabic informally reported to the project officer that the participant information and consent

form included complex research terminology that were difficult for the general community to understand in Arabic. To rectify this, the participant materials were re-translated using a different translation service and reviewed by 5 Arabic-speaking community members.

Potential Uses and Benefits of SMS Text Message Community Panels During COVID-19

The SMS text message community panel is a feasible approach that overcomes many barriers to data collection during a

pandemic. SMS text message community panels allow for noncontact data collection, which is an important attribute during the COVID-19 pandemic, with physical distancing and isolation being key behavioral strategies in preventing COVID-19 spread. SMS text message community panels are potentially able to include people who are more difficult to reach using other data collection methods, such as people from culturally and linguistically diverse backgrounds and differing levels of language proficiency, young people, people who live in rural and remote locations, and people with no fixed address. The method allows for timely data processing, as the data are automatically populated into digital format for analysis, which is particularly pertinent during COVID-19 as situations change quickly. This approach could be used for relatively quick data collection on needs, perceptions, and self-reported behaviors in the context of COVID-19. It is not intended to be a replacement for disease surveillance activities, but represents a potentially important additional method of collecting data.

SMS Text Message Community Panels in Comparison to Other Data Collection Methods

In comparison to online surveys, SMS text message surveys are more specifically tailored to mobile phones. Participants respond to SMS text message survey questions in the same phone app on which they receive the questions, and the app is built into the functionality of mobile phones. By contrast, for online surveys, participants are required to click on a link in an email or social media post, which takes them to a web browser application to complete the survey. While this is a small obstacle, removing any obstacle is beneficial for improving response rates. In addition, SMS text message enables people with mobile phones (not smartphones) to participate.

Online surveys are usually developed as 1 survey with numerous questions. SMS text message community panels send only 1 or 2 questions per week, which means that participants only require a short amount of time to respond to the question(s) each week. In this way, SMS text message community panel surveys are more beneficial for measuring repeated measures at short time intervals. As an example, questions could collect data on how people are feeling during each week, and track changes in relation to situational changes (eg, small outbreaks, changes in lockdown policies, or vaccine rollout announcements).

An important benefit of using SMS text message is that it does not rely on proprietary messaging platforms, such as Facebook or WhatsApp. These proprietary platforms may have privacy or data governance implications. SMS text messages also allow a degree of anonymity as participants can be identified only as their phone number, unlike proprietary platforms, which automatically display names and personal information.

The SMS text message survey method is not appropriate to address all areas of research, particularly those that require more in-depth and detailed inquiry. Closed questions are generally limited to simple ordinal or categorical responses, and responses to open questions are limited to 160 characters before they are split up into multiple messages.

Considerations for Use

Recruitment and Participant Demographics

The experience of the “Shisha No Thanks” project in recruiting SMS text message community panel members demonstrated that recruitment to an already engaged community is effective. Recruitment by texting random mobile numbers with invitations to an SMS text message panel may not be effective. The researchers propose that recruitment to SMS text message community panels should be through channels where people have already established an interest or relationship (eg, signed up to community organization’s database), or through regular recruitment methods (eg, advertisements in newsletters, social media, and personal networks). In addition, this method of data collection is most suitable for recruiting participants from demographic groups who are confident with using SMS text message technology, and frequently use their mobile phone, such as the young people (18-35-year olds) who were the focus of the “Shisha No Thanks” project.

Incentives to Participate

In the “Shisha No Thanks” project, participants reported that reimbursements acted as reminders for people to respond to the survey. Distributing reimbursements using the same platform as was used for the survey questions also facilitated this approach. During COVID-19, easily redeemable reimbursements may be equally important, given the disruptions and other challenges people face.

Data Privacy and Security

In the “Shisha No Thanks” project, data collected were in a nonidentifiable format, and the platform used to create the SMS text message survey used firewall-protected systems and passwords to protect the data. However, SMS text message technology does not use end-to-end message encryption, and so SMS text messages do carry a risk of unauthorized access to the data. Therefore, this method of data collection may not be suitable for sensitive data. Despite the security limitations of SMS text message technology, it is worth noting that studies have found most people do not have privacy or security concerns with using SMS text messages [21,22].

Reducing Barriers

Not everyone will have high levels of literacy, health literacy, or digital literacy. As with all surveys, careful and considered design of the survey questions can help reduce some of these barriers. Enabling the option for people to participate in the “Shisha No Thanks” study in English or Arabic presented a range of challenges including ensuring accurate translation and difficulties in ensuring non-Roman characters displayed correctly when sent via SMS text message. However, the ability to address these challenges illustrates that if inclusiveness is meaningfully considered in design, SMS text message community panels can broaden participation when compared with cross-sectional surveys, as they can be more specifically tailored to the subgroup of interest.

Costs

An important consideration is whether the cost of SMS text message would be a potential barrier for participation. If this method is intended to be used with socioeconomically disadvantaged groups, then this issue needs to be investigated, and it may be important to use toll-free response numbers.

Limitations

This paper presents a case study demonstrating how SMS text message community panels were used for the “Shisha No Thanks” project. Being only 1 case study, there are several limitations in our understanding about how SMS text message community panels could work for other research, particularly in the COVID-19 context. The “Shisha No Thanks” project featured substantial community participation and engagement,

which may have contributed to its ability to recruit and retain participants to the SMS text message community panel. It is unclear how important this initial engagement with the community is to the success of SMS text message community panels. In addition, it is not known how important reimbursements were to the recruitment and retention success of the “Shisha No Thanks” SMS text message community panel. As reimbursement practices vary substantially between studies [23], establishing SMS text message community panels in the future with reduced or no reimbursements would demonstrate whether substantial financial reimbursements are an essential component for this method of data collection. Finally, while we did notice some attrition during the study, further research is needed to identify how long people would be willing to be engaged on such SMS text message community panels.

Acknowledgments

The authors thank Dr Ross MacKenzie for his valuable comments on this paper. LC receives support from the Prevention Research Support Program, funded by the New South Wales Ministry of Health. The “Shisha No Thanks” project was funded by the Cancer Institute NSW.

Conflicts of Interest

None declared.

References

1. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry* 2020 Jun;7(6):547-560 [FREE Full text] [doi: [10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1)] [Medline: [32304649](https://pubmed.ncbi.nlm.nih.gov/32304649/)]
2. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry Res* 2020 May;287:112934 [FREE Full text] [doi: [10.1016/j.psychres.2020.112934](https://doi.org/10.1016/j.psychres.2020.112934)] [Medline: [32229390](https://pubmed.ncbi.nlm.nih.gov/32229390/)]
3. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. *Lancet Public Health* 2020 May;5(5):e256 [FREE Full text] [doi: [10.1016/S2468-2667\(20\)30061-X](https://doi.org/10.1016/S2468-2667(20)30061-X)] [Medline: [32199471](https://pubmed.ncbi.nlm.nih.gov/32199471/)]
4. Liem A, Wang C, Wariyanti Y, Latkin CA, Hall BJ. The neglected health of international migrant workers in the COVID-19 epidemic. *Lancet Psychiatry* 2020 Apr;7(4):e20 [FREE Full text] [doi: [10.1016/S2215-0366\(20\)30076-6](https://doi.org/10.1016/S2215-0366(20)30076-6)] [Medline: [32085842](https://pubmed.ncbi.nlm.nih.gov/32085842/)]
5. Premkumar A, Lovecchio FC, Stepan JG, Kahlenberg CA, Blevins JL, Albert TJ, et al. A Novel Mobile Phone Text Messaging Platform Improves Collection of Patient-Reported Post-operative Pain and Opioid Use Following Orthopedic Surgery. *HSS J* 2019 Feb 22;15(1):37-41 [FREE Full text] [doi: [10.1007/s11420-018-9635-3](https://doi.org/10.1007/s11420-018-9635-3)] [Medline: [30863231](https://pubmed.ncbi.nlm.nih.gov/30863231/)]
6. Lee SSS, Xin X, Lee WP, Sim EJ, Tan B, Bien MPG, et al. The feasibility of using SMS as a health survey tool: an exploratory study in patients with rheumatoid arthritis. *Int J Med Inform* 2013 May;82(5):427-434. [doi: [10.1016/j.ijmedinf.2012.12.003](https://doi.org/10.1016/j.ijmedinf.2012.12.003)] [Medline: [23332386](https://pubmed.ncbi.nlm.nih.gov/23332386/)]
7. Säve-Söderbergh M, Toljander J, Mattisson I, Åkesson A, Simonsson M. Drinking water consumption patterns among adults-SMS as a novel tool for collection of repeated self-reported water consumption. *J Expo Sci Environ Epidemiol* 2018 Mar 14;28(2):131-139 [FREE Full text] [doi: [10.1038/jes.2017.8](https://doi.org/10.1038/jes.2017.8)] [Medline: [28612838](https://pubmed.ncbi.nlm.nih.gov/28612838/)]
8. Kuntsche E, Robert B. Short message service (SMS) technology in alcohol research--a feasibility study. *Alcohol Alcohol* 2009 May 29;44(4):423-428. [doi: [10.1093/alcalc/agg033](https://doi.org/10.1093/alcalc/agg033)] [Medline: [19482879](https://pubmed.ncbi.nlm.nih.gov/19482879/)]
9. Chang T, Gossa W, Sharp A, Rowe Z, Kohatsu L, Cobb EM, et al. Text messaging as a community-based survey tool: a pilot study. *BMC Public Health* 2014 Sep 08;14(1):936 [FREE Full text] [doi: [10.1186/1471-2458-14-936](https://doi.org/10.1186/1471-2458-14-936)] [Medline: [25201051](https://pubmed.ncbi.nlm.nih.gov/25201051/)]
10. Geldsetzer P. Use of Rapid Online Surveys to Assess People's Perceptions During Infectious Disease Outbreaks: A Cross-sectional Survey on COVID-19. *J Med Internet Res* 2020 Apr 02;22(4):e18790 [FREE Full text] [doi: [10.2196/18790](https://doi.org/10.2196/18790)] [Medline: [32240094](https://pubmed.ncbi.nlm.nih.gov/32240094/)]
11. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. *Int J Environ Res Public Health* 2020 Mar 06;17(5):1729 [FREE Full text] [doi: [10.3390/ijerph17051729](https://doi.org/10.3390/ijerph17051729)] [Medline: [32155789](https://pubmed.ncbi.nlm.nih.gov/32155789/)]

12. Kwok KO, Li KK, Chan HHH, Yi YY, Tang A, Wei WI, et al. Community Responses during Early Phase of COVID-19 Epidemic, Hong Kong. *Emerg Infect Dis* 2020 Jul;26(7):1575-1579 [FREE Full text] [doi: [10.3201/eid2607.200500](https://doi.org/10.3201/eid2607.200500)] [Medline: [32298227](https://pubmed.ncbi.nlm.nih.gov/32298227/)]
13. McFadden SM, Malik AA, Aguolu OG, Willebrand KS, Omer SB. Perceptions of the adult US population regarding the novel coronavirus outbreak. *PLoS One* 2020 Apr 17;15(4):e0231808 [FREE Full text] [doi: [10.1371/journal.pone.0231808](https://doi.org/10.1371/journal.pone.0231808)] [Medline: [32302370](https://pubmed.ncbi.nlm.nih.gov/32302370/)]
14. Dos Santos Marques IC, Theiss LM, Johnson CY, McLin E, Ruf BA, Vickers SM, et al. Implementation of virtual focus groups for qualitative data collection in a global pandemic. *Am J Surg* 2021 May;221(5):918-922 [FREE Full text] [doi: [10.1016/j.amjsurg.2020.10.009](https://doi.org/10.1016/j.amjsurg.2020.10.009)] [Medline: [33070983](https://pubmed.ncbi.nlm.nih.gov/33070983/)]
15. Shisha No Thanks! Raising the awareness of the harms of waterpipe smoking in young people from Arabic speaking backgrounds. Centre for Primary Health Care and Equity, UNSW. 2020. URL: <https://cpfce.unsw.edu.au/shisha-no-thanks-raising-awareness-harms-waterpipe-smoking-young-people-arabic-speaking-backgrounds> [accessed 2021-08-08]
16. ShishaNoThanks. URL: <https://www.shishanothanks.org.au/> [accessed 2021-10-14]
17. Cancer Institute NSW. Cancer Institute Tobacco Tracking Survey (CITTS). URL: <https://www.cancer.nsw.gov.au/research-and-data/cancer-data-and-statistics/request-unlinked-unit-record-data-for-research/cancer-institute-tobacco-tracking-survey-citts> [accessed 2021-10-14]
18. Syrian Center for Tobacco Studies, University of Memphis Center for Community Health. Narghile-Waterpipe users Survey. URL: <http://www.scts-sy.org/en/instruments> [accessed 2021-10-14]
19. Yellow Social Media Report 2018. Part One - Consumers. 2018. URL: <https://www.yellow.com.au/social-media-report/> [accessed 2019-02-05]
20. FairWork Ombudsman. Minimum Wages. URL: <https://www.fairwork.gov.au/tools-and-resources/fact-sheets/minimum-workplace-entitlements/minimum-wages> [accessed 2021-08-08]
21. Giguere R, Brown IW, Balán IC, Dolezal C, Ho T, Sheinfil A, et al. Are participants concerned about privacy and security when using short message service to report product adherence in a rectal microbicide trial? *J Am Med Inform Assoc* 2018 Apr 01;25(4):393-400 [FREE Full text] [doi: [10.1093/jamia/ocx081](https://doi.org/10.1093/jamia/ocx081)] [Medline: [29025127](https://pubmed.ncbi.nlm.nih.gov/29025127/)]
22. Dechand S, Naiakshina A, Danilova A, Smith M. In Encryption We Don't Trust: The Effect of End-to-End Encryption to the Masses on User Perception. 2019 Presented at: 2019 IEEE European Symposium on Security and Privacy (EuroS&P); June 17-19, 2019; Stockholm, Sweden p. 401-415. [doi: [10.1109/eurosp.2019.00037](https://doi.org/10.1109/eurosp.2019.00037)]
23. Fry CL, Ritter A, Baldwin S, Bowen KJ, Gardiner P, Holt T, et al. Paying research participants: a study of current practices in Australia. *J Med Ethics* 2005 Sep 01;31(9):542-547 [FREE Full text] [doi: [10.1136/jme.2004.009290](https://doi.org/10.1136/jme.2004.009290)] [Medline: [16131558](https://pubmed.ncbi.nlm.nih.gov/16131558/)]

Edited by G Eysenbach; submitted 18.03.21; peer-reviewed by S Brabyn, A Sudaryanto; comments to author 30.07.21; revised version received 09.08.21; accepted 19.09.21; published 03.11.21.

Please cite as:

Chan L, El-Haddad N, Freeman B, O'Hara BJ, Woodland L, Harris-Roxas B

A Case Study of an SMS Text Message Community Panel Survey and Its Potential for Use During the COVID-19 Pandemic

JMIR Form Res 2021;5(11):e28929

URL: <https://formative.jmir.org/2021/11/e28929>

doi: [10.2196/28929](https://doi.org/10.2196/28929)

PMID: [34612824](https://pubmed.ncbi.nlm.nih.gov/34612824/)

©Lilian Chan, Nouhad El-Haddad, Becky Freeman, Blythe J O'Hara, Lisa Woodland, Ben Harris-Roxas. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.11.2021. 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

Adapting a Mental Health Intervention for Adolescents During the COVID-19 Pandemic: Web-Based Synchronous Focus Group Study

Maria Gabriela Calvo-Valderrama¹, MD; Arturo Marroquín-Rivera¹, MD; Erin Burn², MSc; Laura Ospina-Pinillos¹, MD, PhD; Victoria Bird², PhD; Carlos Gómez-Restrepo¹, MD, MSc, PhD

¹Department of Clinical Epidemiology and Biostatistics, Pontificia Universidad Javeriana, Bogota, Colombia

²Unit for Social and Community Psychiatry, Institute of Population Health Sciences, Queen Mary University of London, London, United Kingdom

Corresponding Author:

Victoria Bird, PhD
Unit for Social and Community Psychiatry
Institute of Population Health Sciences
Queen Mary University of London
Cherry Tree Way
London, E13 8SP
United Kingdom
Phone: 44 20754043802340
Email: v.j.bird@qmul.ac.uk

Abstract

Background: Although focus groups are a valuable qualitative research tool, face-to-face meetings may be difficult to arrange and time consuming. This challenge has been further compounded by the global COVID-19 pandemic and the subsequent lockdown and physical distancing measures implemented, which caused exceptional challenges to human activities. Online focus groups (OFGs) are an example of an alternative strategy and require further study. At present, OFGs have mostly been studied and used in high-income countries, with little information relating to their implementation in low- and middle-income countries (LMICs).

Objective: The aim of this study is to share our experiences of conducting OFGs through a web conferencing service and provide recommendations for future research.

Methods: As part of a broader study, OFGs were developed with adults and adolescents in Colombia during the COVID-19 pandemic. Through a convenience sampling method, we invited eligible participants via email in two different cities of Colombia to participate in OFGs conducted via Microsoft Teams. Researcher notes and discussion were used to capture participant and facilitator experiences, as well as practical considerations.

Results: Technical issues were encountered, but various measures were taken to minimize them, such as using a web conferencing service that was familiar to participants, sending written instructions, and performing a trial meeting prior to the OFG. Adolescent participants, unlike their adult counterparts, were fluent in using web conferencing platforms and did not encounter technical challenges.

Conclusions: OFGs have great potential in research settings, especially during the current and any future public health emergencies. It is important to keep in mind that even with the advantages that they offer, technical issues (ie, internet speed and access to technology) are major obstacles in LMICs. Further research is required and should carefully consider the appropriateness of OFGs in different settings.

(*JMIR Form Res* 2021;5(11):e30293) doi:[10.2196/30293](https://doi.org/10.2196/30293)

KEYWORDS

pandemic; COVID-19; online focus groups; qualitative research; technology; adolescents; public health

Introduction

Focus groups is a commonly used research strategy, particularly within health and social care research, where the experience of individuals, service providers, and the community is vital to innovation and implementation. Focus groups have been an essential tool of qualitative research over the last 50 years [1]; they aim to evaluate different and collective opinions of individuals in a group and pay particular attention to the interaction between participants, thus providing varied information on the topic of interest in a relatively short period [1,2]. Although focus groups offer a valuable method for eliciting a group perspective, their implementation can present a number of challenges. In particular, the need for participants to agree on a mutually convenient time and location can pose difficulties, especially in cases where the participants have existing commitments. Additionally, the recent spread of COVID-19 caused a public health emergency worldwide that affected almost all forms of human activity [3,4]. Physical distancing and lockdown during the pandemic imposed unprecedented challenges to the global population that required innovative strategies to adapt the ways people lived and worked. These recent challenges have also affected scientific research, including qualitative research methods [5], particularly projects that rely on face-to-face data collection [6].

In the past 20 years, the wide use of the internet and the availability of devices such as smartphones have enabled researchers to use web-based platforms to conduct online focus groups (OFGs). This approach has overcome some of the disadvantages of traditional focus groups [7-9]. Owing to the wide range of internet technologies, OFGs can be implemented in various ways. OFGs may be conducted using text-only platforms (ie, chatrooms, discussion boards, and emails) or as virtual group meetings by using other technologies such as webcams, smartphones, and video conferencing services [8]. OFGs can be conducted in two ways—synchronously or asynchronously, depending on whether participants share their opinions simultaneously in the platform or not. Asynchronous OFGs are generally text-based ones, where participants can answer questions through forums, email, or chat in a nonsimultaneous manner. Although the latter may provide a greater sense of anonymity, making it easier for participants to discuss sensitive topics [9,10]. An important disadvantage, however, is that researchers cannot evaluate nonverbal cues, such as eye contact, tone, and body language, that greatly enrich the results of face-to-face focus groups. Furthermore, there is a lack of interaction between participants in these settings [8]. These limitations can be overcome by using webcams in a synchronous focus group [11]. Researchers often experience a closer interaction, similar to the one obtained in a face-to-face setting. Although they tend to have less data production, the quality and level of richness of data is comparable between both settings [7,8].

The use of video conferencing software prevents the need to purchase additional recording devices, overcomes geographical barriers, and can make data transcription easier. However, incorporating new technology in research creates new methodological issues, and for OFGs to function properly, some

requirements must be fulfilled. For instance, they demand participants have a minimum level of digital literacy and a stable internet connection.

These considerations can be especially challenging in low- and middle-income countries (LMICs), where the digital gap is greater as internet access is not universal; resources and infrastructure are scarce; and there is limited funding and little or no support from the government for research activities [2,12-14].

Additionally, it is the responsibility of researchers to carefully select an appropriate web conferencing service to guarantee the privacy and security of participants and the data obtained [2,7,8,15,16]. The latter is of particular significance, since privacy and security breaches have been more frequent during the COVID-19 pandemic, posing serious ethical issues to the conduct of web-based research [17,18].

Research with synchronous OFGs is a growing field; however, most of it has been developed in high-income countries. For example, Kite and Phongsavan [7] compared face-to-face focus groups with OFGs among adults in Australia; they found that OFGs produced rich data similar to face-to-face interactions and that an active discussion between participants was possible even with web-based methods. However, issues with audio, transcription and high levels of participant withdrawal were associated with the web-based modality. To the best of our knowledge, the experience of OFGs has not been reported in LMICs.

Due to the potential of OFGs to overcome the barriers associated with face-to-face methods, especially during times where physical distancing is required, it is important to evaluate their use in populations with different backgrounds. This includes vulnerable individuals and culturally diverse people living in LMICs, where OFGs offer a potentially cost-effective alternative to traditional methods. Consequently, in this paper, we aim to discuss our experiences with conducting OFGs with adults and adolescents during the COVID-19 public health emergency in Colombia.

Methods

We developed OFGs within the framework of the BRiCs study (Building Resilience in Adolescence-Improving Quality of Life for Adolescents With Mental Health Problems in Colombia). This is an ongoing collaborative research project between Queen Mary University of London, United Kingdom, and the Pontificia Universidad Javeriana (PUJ) in Bogotá, the capital of Colombia, and Duitama, an intermediate city, with funding by the UK Research and Innovation (UKRI)—Medical Research Council. This study aims to improve health outcomes for adolescents with depression and anxiety in Colombia by adapting an existing effective app-mediated intervention called DIALOG+ [19-22]. As part of the adaptation component, 10 focus groups were planned in order to collect the end-users' (adolescents and clinicians) and stakeholders' (parents, guardians, youth workers, and educators) opinions, preferences, and information on how to make a resource-oriented intervention (DIALOG+) relevant in this new context and population.

In March 2020, a state of emergency was declared in Colombia due to the COVID-19 pandemic, following which the national government established mandatory quarantine and physical distancing measures. Most cities were placed in lockdown and all travel was restricted. In order to continue the research and to avoid delays concerning project deadlines, an ethics amendment was requested to change the focus group methodology by replacing the 10 face-to-face groups with synchronous OFGs performed through a secure videoconferencing system. Online delivery of the focus groups continued even when lockdown measures were eased due to the fear of contagion and advice to reduce social contact to prevent further disease spread.

Changes to the research protocol were approved by the institutional review board of both academic institutions and clinical settings (protocol FM-CIE-0084-20)

Participants

Participants (adolescents, parents or guardians, clinicians, youth workers or teachers) were recruited from the two clinical settings in Bogotá and Duitama by using a convenience sampling method. To acknowledge the participants for their time, a Col \$40 (approximately US \$12 USD) grocery store voucher was offered to each participant.

Inclusion criteria for the adolescents were (1) age between 13 and 16 years; (2) self-reported current or previous experience of depression and/or anxiety; (3) a willingness to share their experience in an OFG; and (4) capacity to provide informed consent, both by themselves and by a parent or guardian. Parents or guardians were included if they provided care to adolescents aged between 13 and 16 years old with current or previous experience of anxiety and/or depression. Finally, clinicians, educators, and youth workers were (1) required to have experience working with adolescents undergoing depression and/or anxiety and (2) be at least 18 years old.

Data Collection and Analysis

The results of this study focus on the procedures and processes involved in conducting OFGs. This includes describing, in detail, processes such as obtaining informed consent, scheduling meetings, and group facilitation, including any challenges encountered and how they were overcome. Content analysis

results of the OFGs regarding the adaptation of DIALOG+ are beyond the scope of this paper and will be reported separately elsewhere.

Data for the present study were collected through participant observation, with the researchers and group facilitators taking notes during the focus groups. The notes focused on the procedures undertaken, the experience of the group facilitators, and the differences observed between web-based and face-to-face delivery. These observational notes and descriptions of the procedures were gathered by the study coordinator and were discussed during team meetings conducted after the OFG sessions. These reflexive evaluations allowed identification of issues and problems, and action was taken to find potential solutions through discussion with the research team. Additionally, content analysis of the OFG transcriptions was performed when participants expressed their opinions or thoughts related to the methodology used. A final revision of the analyzed content was performed in order to group the information into 7 different categories: consent, booking, facilitation, technical considerations, interaction, and content, which we explore below. This process enabled us to develop themes and guidance, which may be used to guide the future conduct of OFGs.

Results

Below, we first describe the sample and then outline the procedures involved in conducting the OFGs, including any challenges faced.

Sample

A total of 10 OFGs were conducted. In all, 47 participants were approached and only 2 did not participate, which is not unlike the withdrawal rate expected for face-to-face focus groups. One participant did not respond to the invitation email, and the other was unable to participate due to personal circumstances. With a total of 45 participants, each OFG comprised 3 to 7 participants. Participants joined the OFGs mainly from their homes and workplaces. Most of the participants used laptops and desktop computers as their primary device, with a minority using smartphones. Interestingly, tablet devices were not used by the study participants. [Table 1](#) describes the number and general characteristics of participants in each OFG.

Table 1. Participants in each online focus group.

Group	Female (n)	Male (n)
Adolescents		
Bogotá (first)	2	3
Bogotá (second)	3	2
Duitama	3	0
Parents or guardians		
Bogotá	3	0
Duitama	3	0
Clinicians		
Duitama (first)	1	4
Duitama (second)	3	1
Bogotá	5	0
Youth workers or teachers		
Bogotá	4	3
Duitama	5	0

OFG Procedures

Obtaining Informed Consent

The study coordinator invited the participants via email. In the case of parents and adolescents, the invitation email was followed by a phone call. The email and the follow-up phone call provided participants with the relevant information about the study and explained the role of each participant, possible risks, and other information as required. If participants were interested, informed consent was obtained remotely, and individuals were asked to complete a sociodemographic questionnaire. For adolescents, an additional invitation letter explaining the project was sent to their parent or guardian, and we verified that their informed consent had both the adolescent's and guardian's signatures.

As receiving an ink signature for informed consent was challenging due to the circumstances, we obtained an electronic signature from all participants. We also retrieved a signature from the parents or guardians of the participating adolescents. To obtain the electronic signatures, participants were requested to print and sign the informed consent form and send the scanned file to the study coordinator, who was in charge of verifying that every participant had properly filled and sent it prior to each session. All signatures were obtained without difficulty, and none of the participants required assistance or had doubts regarding the process.

Procedure

After obtaining informed consent from the study participants, the study coordinator sent separate invitations for two different meetings. The first meeting included a trial run to check the participant's internet connection and to confirm that all the participants were able to join the OFG and use the videoconferencing software without difficulties. Additionally, during this first trial meeting, the coordinator explained further details of the project and solved logistic and participation

queries. The second invitation was for the OFG session. For both invitations, the time, date, and agenda were included.

Booking

Overall, scheduling the focus groups was not problematic. When sending the invitations to the potential participants, an initial date and time (ie, hour) was stipulated by the study coordinator. We did, however, experience an issue with one OFG session with clinicians—the invitation for a session that was scheduled in the morning (AM) was mistakenly sent for the evening (PM) due to a typographical error. Fortunately, one of the participants double-checked this with the coordinator, and the mistake was rectified in time to enable the rest of participants to join the meeting at the correct time.

Participants' and Researchers' Experience

Facilitation of Focus Groups

Each OFG was facilitated by 2 core members of the research team (LOP and CGR who are psychiatrists and academic researchers) and an anthropologist, who have extensive experience conducting focus groups. Decisions regarding who would facilitate each OFG was based on availability of the team members. Within the group, we allowed multiple people to speak at the same time, to keep the dynamics similar to that of a face-to-face group; however, using the "raising of hands" function on Microsoft Teams was encouraged.

Initially, the facilitator introduced him or herself, provided a general description of the team, and shared the expectations for the session as well as the ground rules (see [Textbox 1](#)). We then asked if every participant had read and understood the informed consent and checked that everyone agreed with recording the session (audio backup recording was also in place). We reminded participants that the audio would be transcribed without any identifying data, so anonymity was ensured, and requested both participants and researchers to activate their web

cameras in order to obtain visual cues. Each participant was then asked to provide a brief introduction.

Facilitators followed session guidelines so that the topics for discussion were consistent in all OFGs. An observer was off camera, taking notes of the visual cues and the process. All

OFGs were conducted within 90 to 120 minutes. In general, the facilitator's role in OFG was more active than in a face-to-face scenario, both for encouraging and moderating participation, as well as maintaining order to avoid simultaneous speaking when a discussion was ongoing.

Textbox 1. Ground rules as per the facilitators' guidelines.

Ground rules:

- Properly introduce yourself and the team. Clarify the purpose of the session and of the data collected.
- Remind participants that the sessions will be audio- and video-recorded and transcribed.
- Explain how confidentiality is ensured.
- Emphasize that there are no right or wrong answers, just points of view and opinions on the DIALOG+ intervention.
- Suggest participants to avoid naming institutions or people when talking about their own experiences, but if they do, remark that it will be erased from the transcription.
- Participants may use a pseudonym.
- Ask participants to speak one at a time to avoid interrupting others.
- Clarify the duration of the session (120 min).
- Remind participants that they can leave or take a break at any point during the discussion.

Technical Considerations

Currently, there are several web conferencing services available to facilitate video calls. To select the best one, we evaluated different options and asked the Information and Communication Technology Service of PUJ for advice. The main criteria included finding a platform that prioritizes security and data privacy. We completed the first 3 OFGs in real time using the web-based platform Cisco Webex [23]. During these first OFGs with clinicians, the initial part of the session was spent resolving issues and concerns related to the platform, such as how to join the meeting and activation of the camera and audio, which made communication slower. Participants mentioned that they felt the video conferencing was not as easy as they expected because they were not familiar with the interface, which is not widely used within Colombia. Therefore, we decided to host the remaining OFGs with Microsoft Teams [24], a technology that is more commonly used within Colombia.

In order to overcome this challenge and familiarize participants with the platform, we sent an instructions manual, via email, explaining how to set up a Microsoft Teams account and join the virtual meeting. To further assist the participants, we scheduled a short meeting with the study coordinator before each OFG, to test connectivity and solve technical issues. As suggested by Kite and Phongsavan [7], we also encouraged early login to the platform on the day of the session [7]. Nonetheless, most participants joined the meeting a few minutes after the stipulated time.

When using Microsoft Teams, we noticed that joining the meeting was easier for those participants who had the desktop app installed on their computers than for those using the browser version. We therefore recommended installing the app prior to the OFG. However, we did experience issues with the audio during one OFG session. After the participant changed his microphone and restarted the software on his computer without

success, we suggested using the browser version that solved the issue.

In general, all of these measures undertaken helped us to utilize most of the scheduled time for each session with discussion relevant to the research, rather than with technical discussion; it also made participants more involved, even if personal matters, including children, pets, and phones, would sometimes distract them momentarily during participation from their homes.

As researchers, organizing conventional face-to-face focus groups is challenging, especially in large urban areas such as Bogotá; hence, OFGs were perceived as a good alternative. In a smaller city such as Duitama, on the other hand, we had concerns such as less stable internet connection. However, network coverage was better than anticipated. Therefore, this dismissed our concern regarding internet connectivity and device availability.

As expected with adolescents (both in Bogotá and Duitama), we did not face any technical challenges, and participants' use of the web conferencing service was seamless. In the test meeting, we had no discussion about the use of the platform, and only general concerns about the informed consent form and sociodemographic questionnaire were addressed.

Adult participants in Duitama seemed to face more challenges with using the web conferencing service, and they had more queries about the platform, which needed to be resolved. It is possible that individuals in Duitama, which is an intermediate city, were not as familiar with conferencing services as those within major urban areas such as Bogotá. Since both settings were urban areas, we did not encounter problems with internet connection. Moreover, other challenges, such as those pertaining to audio quality, as reported previously [7,25], were not a major issue in our study, which made audio transcription easy. All sessions were recorded using web conferencing recording

features, and a backup audio recording was made by the study coordinator using the computer's audio recorder.

Participation Interaction

All participants agreed to activate their cameras, this enabled us to ensure similar interactions to that expected in a face-to-face focus group. Questions were presented following a predetermined order, from general to more specific topics according to a previously developed facilitator's guide. Raising hand via emoji within the software or physically by the participants raising their hand via the camera allowed participants to take turns. Although sometimes participants tended to speak simultaneously making it difficult to hear all opinions, it was the duty of the facilitators to remind individuals to take turns as would happen in a traditional face-to-face group.

The facilitator encouraged all participants to share their opinions on each topic presented. Participation was modest during the initial parts of every OFG, which can also be common in face-to-face focus groups [26]. We noticed that after a few questions, engagement increased, and participation began to be spontaneous.

This modality lacked features such as small chats and paired discussions between participants. It was obvious that participants who knew each other before the OFG (eg, clinicians and teachers and youth workers) were more engaged and participated more than those who did not know each other beforehand. Spontaneous social interaction and acquaintance between unfamiliar participants did not happen during the groups. Given that the online context can overlook some nonverbal cues, the facilitators had to rely on asking direct questions to invite people to share opinions and make sure that everyone could share their view without interruption.

Overall, the feedback obtained suggested that both researchers and participants perceived OFGs as a good alternative to face-to-face groups. Particular logistical advantages were discussed. In both settings, particularly in Bogotá, OFGs were perceived as less time consuming because there was no need to factor in travel time, which in a large city or rural area can be significant.

Additionally, costs were diminished because we did not have to consider transportation fees, hospitality, or additional recording equipment because all sessions were recorded through the recording features of the web conferencing platform.

Content of the Focus Groups

Our project discussed an app (DIALOG+) aimed to improve outcomes of depression and anxiety in adolescents based in Colombia. As mentioned, participating adolescents had self-reported current or previous experience with anxiety and/or depression, and all stakeholders had experience in this field. Discussing mental health has the potential to open up sensitive topics that can trigger distressing responses from participants, particularly adolescents, who may require additional support. We did not experience this issue in our study, as none of the participants reported feeling distressed by the topics discussed.

However, we consider that for all focus groups, regardless of the modality (ie, face-to-face or web-based focus groups), this aspect must be considered when sensitive topics are discussed as part of the study. Strategies to manage participant distress should be discussed between researchers, such as providing additional resources (eg, helpline and crisis contacts), clinical staff on site, or appropriate referral pathways.

In our case, 2 of the facilitators (CG and LOP) were also clinicians with extensive experience in child and adolescent mental health, and they were available to be contacted during or after the session if a participant required help, in which case they would evaluate the need for treatment or additional interventions and refer them to the appropriate services.

Discussion

During unprecedented circumstances, OFGs were a useful tool to guarantee research continuity in cases where physical distancing was mandatory [27]. This change of methodology generated new knowledge and skills for the members of our research team and enabled us to reduce significant delays in meeting our research deadlines as well as collect considerable data that allowed us to fulfill the aim of this phase of our research project. Furthermore, the ease of organizing and scheduling OFGs, especially for clinicians and individuals who lived in larger geographical areas, offered a viable alternative to face-to-face meetings. However, with the introduction of a new technology to a traditional research method, we expected new challenges. Based on our experience conducting OFGs with multiple stakeholders, a summary of our recommendations on performing OFGs is shown in [Textbox 2](#).

Textbox 2. Summary of recommendations to perform OFG.**Before the online focus group:**

- Verify dates and hours scheduled. Double-check for any typographical errors or autocorrections in invitations.
- Consider several web conferencing services. Consider privacy and security settings, as well as familiarity of participants according to local contexts.
- Schedule a trial meeting of short duration before the actual online focus group session to check participant's internet connection and solve any doubts.
- Send a brief instructive email explaining how to join the meeting, activate audio and camera, and install the selected platform desktop version.
- Consider incentives to minimize possible withdrawals.

During the online focus group:

- Encourage early login and use of headphones with built in microphone, if available.
- Schedule more time than you would for a face-to-face focus group. This allows facilitators to perform ice breaker activities and deal with technical issues that may arise.
- Fewer participants (between 4 and 6) may have more interaction and active participation than larger groups.

The majority of challenges faced were related to technology literacy and a lack of familiarity with new videoconferencing software. While selecting the platform to conduct OFGs, it is important to consider which web conferencing platform is the most suitable for the particular target population [7,8]. We suggest considering the security and privacy settings offered by each one, as well as the familiarity participants might have with them [11]. Research suggests that participants tend to be less distracted by the web-based platform when they are more familiar with it [15]. We noticed less discussion regarding technical issues when we used Microsoft Teams compared to Cisco Webex, maybe because the former is more commonly used in Colombia. Therefore, when considering which platform to use for OFGs, it is important to consider the preference and familiarity of the target population.

To increase participant familiarity with the videoconferencing system, particularly for those that had not used it before, sending instructions on how to join meetings and setting up the microphone and web camera can be useful, as well as prompting individuals to install the desktop app. For the latter, verifying that the application can access the microphone and camera is vital (especially on Windows PCs). Prior to conducting a focus group (both face-to-face and OFGs), we recommend participants to double-check dates and hours on invitations, pay special attention to any typographical errors, or autocorrections that might have occurred.

Another strategy that helped us to use the time of the OFG session for research-related topics was performing a trial meeting the day before to solve concerns (related to the web platform or the research itself), provide more personalized assistance, and encourage participants to login early on the day of the session. Despite specific instructions and the test trial, minor technical issues still occurred. However, overall, these measures considerably minimized technical and procedural issues. Since technical challenges cannot be prevented in their totality, we consider that it is better to schedule more time for an OFG than you would for a face-to-face focus group [7]. Allowing additional time may also help because the novelty of the modality can cause participants to be initially apprehensive

about participating in the discussion. Therefore, additional time can be used to implement strategies that stimulate contribution from each participant, such as introductory and ice-breaker activities.

Since both settings of our study (Bogotá and Duitama) were urban, access to computers and a reliable internet connection were available and made it easier for us to conduct the research. If we were to conduct OFGs in rural areas, it is likely that we would have faced greater challenges, especially in an LMIC like Colombia.

Unlike the experiences reported by Tuttas [25] and Kite and Phongsavan [7], we did not face any challenges regarding sound quality. Most participants were using headphones with built-in microphones, so we were able to allow all participants to be unmuted, thus allowing a more fluent discussion without much background noise. This also meant that we did not face any problems with the audio transcription.

We paid particular attention to the aspect of informed consent, especially since the study involved adolescent participants. All participants received the informed consent forms electronically and were able to seamlessly send a scanned version with their signatures, as well as their parent's or guardians' signature, as required. This avoided any form of face-to-face contact with the participants during the COVID-19 pandemic.

Another factor that must be kept in mind, is the number or participants that is optimal for OFGs. In our experience, we consider that the optimum number of participants for OFGs would be between 3 and 5. This number is fewer than would be expected in a traditional face-to-face focus group [28,29]. Small groups were preferred as they facilitated more interaction. When the number of participants was higher, we observed that some participants answered most questions and the rest would rarely speak (unless directly asked), or the group interactions were less extensive.

We consider that the skills of the facilitator are crucial for directing OFGs. Certain nonverbal cues are visible via the webcam (eg, head nodding or raising of hands), but others, such

as direct eye contact that can signal eagerness to speak, are lost. Therefore, interaction has to be more direct. The facilitator has to actively invite people to communicate and has to be tactful when there are multiple interactions, allowing everyone to express their opinion without interrupting abruptly. With this in mind, explaining ground rules on participation at the beginning of the session is of great importance.

Previous studies have shown higher rates of dropouts [11,15] and withdrawal in OFG, especially when there is a discussion of sensitive topics [25]. We did not experience this, probably due to the nature of the topic discussed, and the incentive of a receiving store voucher could also have minimized dropouts. Considering these kinds of incentives might help reduce the rates of withdrawal. Another factor that could have contributed is that these OFGs were conducted during lockdown and the consequent physical distancing measures. Therefore, most of our participants were working or studying from home and probably had more availability to participate than in another setting.

It is important to highlight an aspect of our research, which is developing OFGs with adolescents. The use of the internet in this age group can be seen as a less intimidating way of encouraging participation [9], and their familiarity with the online world gives them an augmented sense of control. OFGs were perceived as a less hierarchical interaction than the one in a face-to-face contact [30]. In our case, adolescents were clearly fluent in the use of the web conferencing software, and we noticed that participants, across all groups, actively participated in both study sites.

Limitations

The major limitation of this study was that given the current situation, the quality of the information obtained with the OFGs could not be compared to that obtained from a face-to-face group. This means that we do not know if the quality of data

was significantly different. When using a familiar web conferencing service, participants mentioned feeling comfortable with the methodology and stated that they expressed their opinions without feeling restricted by the form of communication. Therefore, we consider that the quality of information obtained would have been similar had we conducted a face-to-face focus group. We believe that sharing our experience of conducting OFGs with adolescents in an LMIC, in the context of the COVID-19 crisis, can be valuable for future researchers.

Conclusions

Overall, our experience using web conferencing services to perform OFGs was successful.

We consider that the current technological advancements provide OFGs with great potential in research settings, especially in the current global pandemic that has made it difficult to conduct research. A positive aspect of the current pandemic may be that a greater number of people who were unlikely to use web conferencing services are now familiar with them.

It is important to note that even if this modality can overcome geographical barriers, technical issues such as internet speed and access to equipment are great obstacles in LMICs, especially in rural areas. The access and knowledge to these platforms reflect a level of access to technological resources that is not yet universal, which means that many groups can be underrepresented [2]. For example, certain groups, such as the elderly or those from lower socioeconomic backgrounds, may lack access to technology and/or technical competencies required.

For the purposes of our research, OFGs allowed continuity with satisfactory results, and the objectives for the initial stage of our study were met thanks to the quality of data obtained. Further testing of this method is required to overcome current limitations.

Conflicts of Interest

None declared.

References

1. Parker A, Tritter J. Focus group method and methodology: current practice and recent debate. *International Journal of Research & Method in Education* 2006 Apr;29(1):23-37. [doi: [10.1080/01406720500537304](https://doi.org/10.1080/01406720500537304)]
2. Gaiser TJ. Conducting on-line focus groups. *Soc Sci Comput Rev* 2016 Aug 19;15(2):135-144. [doi: [10.1177/089443939701500202](https://doi.org/10.1177/089443939701500202)]
3. Walker PGT, Whittaker C, Watson OJ, Baguelin M, Winskill P, Hamlet A, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* 2020 Jul 24;369(6502):413-422 [FREE Full text] [doi: [10.1126/science.abc0035](https://doi.org/10.1126/science.abc0035)] [Medline: [32532802](https://pubmed.ncbi.nlm.nih.gov/32532802/)]
4. Baldwin R, Weder di Mauro B, editors. *Economics in the time of COVID-19*. London, UK: Centre for Economic Policy Research; 2020.
5. Bikson M, Hanlon CA, Woods AJ, Gillick BT, Charvet L, Lamm C, et al. Guidelines for TMS/tES clinical services and research through the COVID-19 pandemic. *Brain Stimul* 2020;13(4):1124-1149 [FREE Full text] [doi: [10.1016/j.brs.2020.05.010](https://doi.org/10.1016/j.brs.2020.05.010)] [Medline: [32413554](https://pubmed.ncbi.nlm.nih.gov/32413554/)]
6. Marhefka S, Lockhart E, Turner D. Achieve research continuity during social distancing by rapidly implementing individual and group videoconferencing with participants: key considerations, best practices, and protocols. *AIDS Behav* 2020 Jul;24(7):1983-1989 [FREE Full text] [doi: [10.1007/s10461-020-02837-x](https://doi.org/10.1007/s10461-020-02837-x)] [Medline: [32240428](https://pubmed.ncbi.nlm.nih.gov/32240428/)]

7. Kite J, Phongsavan P. Insights for conducting real-time focus groups online using a web conferencing service. *F1000Res* 2017;6:122 [[FREE Full text](#)] [doi: [10.12688/f1000research.10427.1](https://doi.org/10.12688/f1000research.10427.1)] [Medline: [28781752](https://pubmed.ncbi.nlm.nih.gov/28781752/)]
8. Abrams KM, Wang Z, Song YJ, Galindo-Gonzalez S. Data richness trade-offs between face-to-face, online audiovisual, and online text-only focus groups. *Soc Sci Comput Rev* 2014 Jun 03;33(1):80-96. [doi: [10.1177/0894439313519733](https://doi.org/10.1177/0894439313519733)]
9. Fox FE, Morris M, Rumsey N. Doing synchronous online focus groups with young people: methodological reflections. *Qual Health Res* 2007 Apr;17(4):539-547. [doi: [10.1177/1049732306298754](https://doi.org/10.1177/1049732306298754)] [Medline: [17416707](https://pubmed.ncbi.nlm.nih.gov/17416707/)]
10. Wettergren L, Eriksson LE, Nilsson J, Jervaeus A, Lampic C. Online focus group discussion is a valid and feasible mode when investigating sensitive topics among young persons with a cancer experience. *JMIR Res Protoc* 2016 May 09;5(2):e86 [[FREE Full text](#)] [doi: [10.2196/resprot.5616](https://doi.org/10.2196/resprot.5616)] [Medline: [27161146](https://pubmed.ncbi.nlm.nih.gov/27161146/)]
11. Daniels N, Gillen P, Casson K, Wilson I. STEER: factors to consider when designing online focus groups using audiovisual technology in health research. *Int J Qual Methods* 2019 Nov 19;18:160940691988578. [doi: [10.1177/1609406919885786](https://doi.org/10.1177/1609406919885786)]
12. Lamprea E, García J. Closing the gap between formal and material health care coverage in Colombia. *Health Hum Rights* 2016 Dec;18(2):49-65 [[FREE Full text](#)] [Medline: [28559676](https://pubmed.ncbi.nlm.nih.gov/28559676/)]
13. Rubio MA, Dennis R, Dominguez MT, Suarez R, Olarte-Sierra MF, Palacio D, et al. Challenges to the improvement of Colombian medical brigades aimed at the diagnosis of congenital heart disease: A qualitative approach. *Glob Public Health* 2019 Aug;14(8):1193-1203. [doi: [10.1080/17441692.2018.1549266](https://doi.org/10.1080/17441692.2018.1549266)] [Medline: [30468098](https://pubmed.ncbi.nlm.nih.gov/30468098/)]
14. Waithira N, Mutinda B, Cheah PY. Data management and sharing policy: the first step towards promoting data sharing. *BMC Med* 2019 Apr 17;17(1):80 [[FREE Full text](#)] [doi: [10.1186/s12916-019-1315-8](https://doi.org/10.1186/s12916-019-1315-8)] [Medline: [30992010](https://pubmed.ncbi.nlm.nih.gov/30992010/)]
15. Fielding N, Lee RM, Blank G, editors. *The SAGE Handbook of Online Research Methods*. Thousand Oaks, CA: Sage Publications; Mar 2017:684.
16. Birnbaum MH. Human research and data collection via the internet. *Annu Rev Psychol* 2004;55:803-832. [doi: [10.1146/annurev.psych.55.090902.141601](https://doi.org/10.1146/annurev.psych.55.090902.141601)] [Medline: [14744235](https://pubmed.ncbi.nlm.nih.gov/14744235/)]
17. Kagan D, Alpert G, Fire M. Zooming into video conferencing privacy and security threats. arXiv. Preprint posted online on July 2, 2020 [[FREE Full text](#)]
18. Aiken A. Zooming in on privacy concerns: Video app Zoom is surging in popularity. In our rush to stay connected, we need to make security checks and not reveal more than we think. *Index on Censorship* 2020 Jun 17;49(2):24-27 [[FREE Full text](#)] [doi: [10.1177/0306422020935792](https://doi.org/10.1177/0306422020935792)]
19. Priebe S, Kelley L, Omer S, Golden E, Walsh S, Khanom H, et al. The effectiveness of a patient-centred assessment with a solution-focused approach (DIALOG+) for patients with psychosis: a pragmatic cluster-randomised controlled trial in community care. *Psychother Psychosom* 2015;84(5):304-313 [[FREE Full text](#)] [doi: [10.1159/000430991](https://doi.org/10.1159/000430991)] [Medline: [26278784](https://pubmed.ncbi.nlm.nih.gov/26278784/)]
20. Priebe S, Golden E, Kingdon D, Omer S, Walsh S, Katevas K, et al. Effective patient-clinician interaction to improve treatment outcomes for patients with psychosis: a mixed-methods design. *Programme Grants Appl Res* 2017 Feb;5(6):1-160. [doi: [10.3310/pgfar05060](https://doi.org/10.3310/pgfar05060)] [Medline: [28252893](https://pubmed.ncbi.nlm.nih.gov/28252893/)]
21. Priebe S, Kelley L, Golden E, McCrone P, Kingdon D, Rutterford C, et al. Effectiveness of structured patient-clinician communication with a solution focused approach (DIALOG+) in community treatment of patients with psychosis--a cluster randomised controlled trial. *BMC Psychiatry* 2013 Jun 26;13:173 [[FREE Full text](#)] [doi: [10.1186/1471-244X-13-173](https://doi.org/10.1186/1471-244X-13-173)] [Medline: [23802602](https://pubmed.ncbi.nlm.nih.gov/23802602/)]
22. Priebe S, Fung C, Sajun SZ, Alinaitwe R, Giacco D, Gómez-Restrepo C, et al. Resource-oriented interventions for patients with severe mental illnesses in low- and middle-income countries: trials in Bosnia-Herzegovina, Colombia and Uganda. *BMC Psychiatry* 2019 Jun 14;19(1):181 [[FREE Full text](#)] [doi: [10.1186/s12888-019-2148-x](https://doi.org/10.1186/s12888-019-2148-x)] [Medline: [31200671](https://pubmed.ncbi.nlm.nih.gov/31200671/)]
23. Cisco Webex. URL: <https://www.webex.com/> [accessed 2021-10-25]
24. Microsoft Teams. URL: <https://www.microsoft.com/en/microsoft-365/microsoft-teams/group-chat-software> [accessed 2021-10-25]
25. Tuttas CA. Lessons learned using web conference technology for online focus group interviews. *Qual Health Res* 2015 Jan;25(1):122-133. [doi: [10.1177/1049732314549602](https://doi.org/10.1177/1049732314549602)] [Medline: [25192765](https://pubmed.ncbi.nlm.nih.gov/25192765/)]
26. Krueger R, Casey M. *Designing and Conducting Focus Group Interviews*. 2002 Oct. URL: <https://www.eiu.edu/ihec/Krueger-FocusGroupInterviews.pdf> [accessed 2021-10-25]
27. Mesa Vieira C, Franco OH, Gómez Restrepo C, Abel T. COVID-19: The forgotten priorities of the pandemic. *Maturitas* 2020 Jun;136:38-41 [[FREE Full text](#)] [doi: [10.1016/j.maturitas.2020.04.004](https://doi.org/10.1016/j.maturitas.2020.04.004)] [Medline: [32386664](https://pubmed.ncbi.nlm.nih.gov/32386664/)]
28. Kitzinger J. The methodology of focus groups: the importance of interaction between research participants. *Sociol Health & Illness* 1994 Jan;16(1):103-121. [doi: [10.1111/1467-9566.ep11347023](https://doi.org/10.1111/1467-9566.ep11347023)]
29. Ospina-Pinillos L, Davenport T, Mendoza Diaz A, Navarro-Mancilla A, Scott EM, Hickie IB. Using participatory design methodologies to co-design and culturally adapt the Spanish version of the mental health eClinic: qualitative study. *J Med Internet Res* 2019 Aug 02;21(8):e14127 [[FREE Full text](#)] [doi: [10.2196/14127](https://doi.org/10.2196/14127)] [Medline: [31376271](https://pubmed.ncbi.nlm.nih.gov/31376271/)]
30. Kitchin R. *Cyberspace: The World in the Wires*. New York, NY: John Wiley & Sons, Inc; Jun 1998:256.

Abbreviations

BRiCs: Building Resilience in Adolescence-Improving Quality of Life for Adolescents With Mental Health Problems in Colombia

LMICs: low- and middle-income countries

OFG: online focus group

PUJ: Pontificia Universidad Javeriana

UKRI: UK Research and Innovation

Edited by G Eysenbach; submitted 10.05.21; peer-reviewed by N Daniels, I Ghodratiostani; comments to author 04.08.21; revised version received 13.09.21; accepted 20.09.21; published 03.11.21.

Please cite as:

Calvo-Valderrama MG, Marroquín-Rivera A, Burn E, Ospina-Pinillos L, Bird V, Gómez-Restrepo C

Adapting a Mental Health Intervention for Adolescents During the COVID-19 Pandemic: Web-Based Synchronous Focus Group Study

JMIR Form Res 2021;5(11):e30293

URL: <https://formative.jmir.org/2021/11/e30293>

doi: [10.2196/30293](https://doi.org/10.2196/30293)

PMID: [34637395](https://pubmed.ncbi.nlm.nih.gov/34637395/)

©Maria Gabriela Calvo-Valderrama, Arturo Marroquín-Rivera, Erin Burn, Laura Ospina-Pinillos, Victoria Bird, Carlos Gómez-Restrepo. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.11.2021. 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

COVID-19 Knowledge, Attitudes, and Practices Among People in Bangladesh: Telephone-Based Cross-sectional Survey

Md Golam Rabbani¹, MSS; Orin Akter², MSS; Md Zahid Hasan², MSS; Nandeeta Samad³, MSc; Shehrin Shaila Mahmood², PhD; Taufique Joarder¹, PhD

¹Public Health Foundation, Bangladesh, Dhaka, Bangladesh

²Health Systems and Population Studies Division, International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh

³Department of Public Health, North South University, Dhaka, Bangladesh

Corresponding Author:

Md Golam Rabbani, MSS

Public Health Foundation, Bangladesh

54 Inner Circular Road

Scaut Market, Naya Paltan

Dhaka, 1217

Bangladesh

Phone: 880 1732292488

Email: rabbaniduihe@gmail.com

Abstract

Background: The world has been grappling with the COVID-19 pandemic, a dire public health crisis, since December 2019. Preventive and control measures have been adopted to reduce the spread of COVID-19. To date, the public's knowledge, attitudes, and practices regarding COVID-19 across Bangladesh have been poorly understood. Therefore, it is important to assess people's knowledge, attitudes, and practices (KAP) toward the disease and suggest appropriate strategies to combat COVID-19 effectively.

Objective: This study aimed to assess the KAP of Bangladeshi people toward COVID-19 and to identify their determinants.

Methods: We conducted a country-wide cross-sectional telephonic survey from May 7 to 29, 2020. A purposive sampling method was applied, and adult Bangladeshi citizens who have mobile phones were approached to participate in the survey. Interviews were conducted based on verbal consent. Multiple logistic regression analyses and several tests were performed to identify the factors associated with KAP related to COVID-19.

Results: A total of 492 of 576 Bangladeshi adults aged 18 years and above completed the interview, with a response rate of 85.4% (492/576). Of the 492 participants, 321 (65.2%) were male, and 304 (61.8%) lived in a rural area. Mean scores for knowledge, attitudes, and practices were 10.56 (SD 2.86), 1.24 (SD 0.83), and 3.17 (SD 1.5), respectively. Among the 492 respondents, 273 (55.5%) had poor knowledge, and 251 (49%) expressed a negative attitude; 192 out of 359 respondents (53.5%) had poor practices toward COVID-19. Mean scores of knowledge, attitudes, and practices differed significantly across various demographic and socioeconomic groups. Rural residents had lower mean scores of knowledge (mean 9.8, SD 3.1, $P<.001$) and adherence to appropriate practice measures (mean 4, SD 1.4, $P<.001$) compared to their urban counterparts. Positive and statistically strong correlations between knowledge and attitudes ($r=0.21$, $P<.001$), knowledge and practices ($r=0.45$, $P<.001$), and attitudes and practices ($r=0.27$, $P<.001$) were observed. Television (53.7%) was identified as the major source of knowledge regarding COVID-19. Almost three-quarters of the respondents (359/492, 73%) went outside the home during the lockdown period. Furthermore, the study found that good knowledge (odds ratio [OR] 3.13, 95% CI 2.03-4.83, and adjusted OR 2.33, 95% CI 1.16-4.68) and a positive attitude (OR 2.43, 95% CI 1.59-3.72, and adjusted OR 3.87, 95% CI 1.95-7.68) are significantly associated with better practice of COVID-19 health measures.

Conclusions: Evidence-informed and context-specific risk communication and community engagement, and a social and behavior change communication strategy against COVID-19 should be developed in Bangladesh based on the findings of this study, targeting different socioeconomic groups.

(JMIR Form Res 2021;5(11):e28344) doi:[10.2196/28344](https://doi.org/10.2196/28344)

KEYWORDS

COVID-19; knowledge; attitude; practice; risk communication and community engagement; social and behavior change communication; Bangladesh; COVID-19; risk; pandemic; risk communication

Introduction

Since the end of 2019, the world has been confronting the continuous morbidity and mortality caused by COVID-19 [1,2]. Bangladesh is one of the most affected countries, and it is currently experiencing a new surge of infections. As of March 25, 2021, there had been reports of over 125 million confirmed cases of COVID-19, including more than 2.75 million deaths globally [3]. As of March 22, 2021, Bangladesh was among the top 34 countries, with 0.47% of the total global COVID-19 cases [4].

In Bangladesh, the first confirmed COVID-19 cases and death were reported on March 8 and 18, 2020, respectively [5]. Since then, the pandemic has spread gradually over the whole country, and the number of affected people has been increasing [4]. It has been presumed that the consequences of COVID-19 will be devastating in Bangladesh as, with a population of approximately 165 million, it is the second most densely populated country in the world, with approximately 1240 people per square kilometer [6-8]. Moreover, a large proportion of the population still lives below the poverty line, and almost half of the population is exposed to multiple socioeconomic vulnerabilities [6,9]. Further, Bangladesh suffers from a low literacy rate [7], potentially exposing the population to unfavorable knowledge, attitudes, and practices (KAP) toward a persisting pandemic. Evidence has shown that socioeconomically vulnerable communities lag in access to health, education, and information, which may aggravate the spread of an infectious disease [10,11].

In a short time, after the identification of the first case, Bangladesh was recognized as the second most affected country in South Asia. To control the spread of COVID-19, therefore, the government declared a nationwide “general holiday” (lockdown) for a few months and recommended that its people maintain the standard guidelines [12-14]. Since the beginning of the pandemic, control of the spread of SARS-CoV-2 has been considered the most effective measure to protect people. Thus, regarding preventing the spread of COVID-19, nonclinical interventions such as restricting unnecessary movement, maintaining social distance, wearing masks, and frequent and appropriate handwashing have been suggested by the World Health Organization (WHO) [15-17]. Maintaining such interventions largely depends on people’s KAP, as existing scientific evidence has shown that sociodemographic and economic characteristics are associated with the level of people’s knowledge and attitudes, which safeguard against disease spread [10,18].

Inadequate knowledge, negative attitudes, and poor practices toward an infectious disease may lead to superfluous threats, rumors, confusion, and even unnecessary fear, which may aggravate the pandemic. Misconception, contradictory

perception, and practice of unscientific remedies for COVID-19 were found among Bangladeshi people. Thus, it is important to understand the KAP regarding COVID-19 among the general population to prevent the spread of the infection. This study will contribute to filling the gap that currently exists. The objective of this study was to assess the KAP toward COVID-19 of Bangladeshi residents. The findings from this study are expected to help policy makers design context-specific social and behavior change strategies for the population in Bangladesh and in other similar settings.

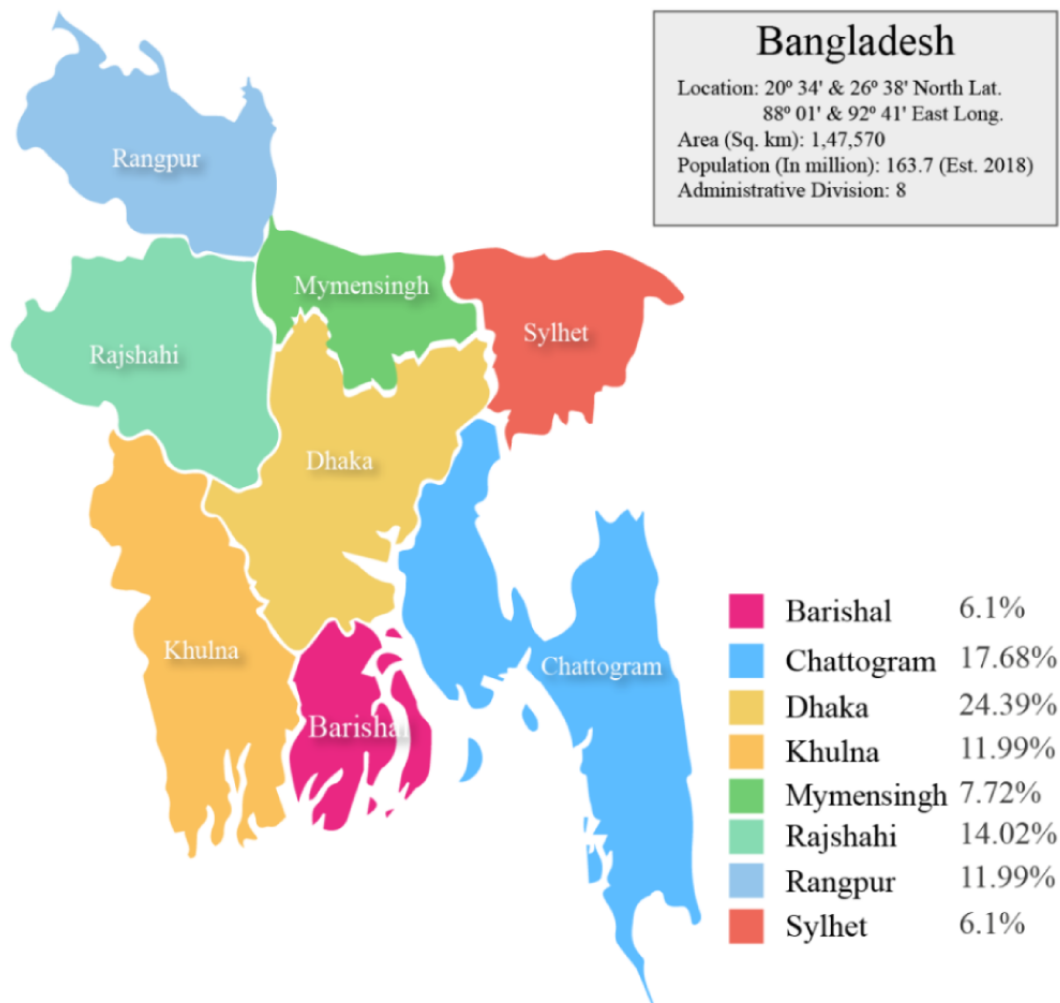
Methods

Study Design and Setting

A cross-sectional telephone survey was conducted applying a quantitative approach. The survey was conducted from May 7 to 29, 2020, during the lockdown period in Bangladesh. The study was conducted in all 8 administrative divisions (Barishal, Chattogram, Dhaka, Khulna, Mymensingh, Rajshahi, Rangpur, and Sylhet) of the country.

Recruitment Procedure

A purposively selected sample of individuals was interviewed to measure the KAP toward COVID-19. The inclusion criteria included Bangladeshi citizens with a minimum age of 18 years who are in possession of a mobile phone. The researchers collected contact numbers of the study population through direct contact or via their network, consisting of family members, friends, and relatives, regardless of their demographic and socioeconomic status according to the inclusion criteria. To ensure the fulfillment of the inclusion criteria, we asked the respondents about their birth year and whether they had a national identity card (which is only available to Bangladeshis over the age of 18 years). When collecting contact numbers, the people were informed that they might be contacted for an interview. With this method, we developed a pool of eligible and initially interested respondents. Before data collection, we estimated the sample size. We assumed that 50% ($P=.05$) of the population has the proper KAP toward COVID-19. With a 95% confidence level and 5% precision level, a representative sample of 384 was estimated using this proportion. The sample size increased to 460 and 576 after adjusting for the 1.2 design effect and the 20% nonresponse rate. We approached 576 individuals who were interested in taking part in this study from 8 divisions proportionately to the population of the respective divisions (Barishal 5.7%, Chattogram 17.5%, Dhaka 23.3%, Khulna 11.9%, Mymensingh 7.4%, Rajshahi 14.3%, Rangpur 11.8%, and Sylhet 6%) [19] (Figure 1), of whom 504 individuals agreed to participate in the interview and another 72 individuals were not interested in participating in the interview. A total of 492 completed interviews were used to analyze data after excluding incomplete interviews.

Figure 1. Proportion of respondents who completed interviews in 8 administrative divisions of Bangladesh.

Study Instrument and KAP Measure

We developed a questionnaire ([Multimedia Appendix 1](#)) by following the national guidelines for clinical and community management of COVID-19 introduced by the Government of Bangladesh, WHO reports, and literature [15,20-22]. The questionnaire was then translated into Bengali, contextualized for Bangladesh, reviewed by 2 sectoral experts, piloted in 15 adult people, and finalized to ensure the respondents understood all the KAP-related questions. The individual survey questionnaire consisted of two sections. The first section focused on sociodemographic and economic characteristics, including age, gender, education, occupation, current residence, religion, marital status, number of household members, living conditions, toilets in the respondent's current residence, and income level. The second section covered KAP-related questions (including 14 items for knowledge, 2 items for attitudes, and 6 items for practices), and the overall Cronbach α of KAP questions regarding the final data was .86, which indicates a satisfactory level of internal consistency [23].

The knowledge questions focused on clinical presentations of COVID-19, transmission routes, prevention, and infection control. The responses to these questions included "yes" and "no" options, with an additional "don't know" option. A correct answer was assigned a value of 1, and an incorrect or "don't

know" answer was assigned a value of 0. The total knowledge score ranged from 0 (no correct answers) to 14 (all correct answers), with a higher score indicating better knowledge of COVID-19. To determine the knowledge level, following Bloom's cutoff point [24], we determined that 80% to 100% positive responses indicated good knowledge and percentages of positive responses below that indicated poor knowledge.

A similar scoring approach was used for classifying positive and negative attitudes and for good and poor practices. To measure attitudes toward COVID-19, selected participants were asked 2 questions: whether they agreed, disagreed, or were not sure that the pandemic would be successfully controlled across the world, and their level of confidence that Bangladesh would win the battle against COVID-19.

Similarly, to measure practices, 5 self-reported statements were used, namely, whether respondents had avoided crowded places while being outside; maintained the recommended social distance of 1 m (or approximately 3 feet); worn a mask regularly and correctly when outside; washed their hands regularly and thoroughly after coming from outside and before eating or touching the mouth, nose, or eyes; and maintained the recommended hand washing duration of 20 to 30 seconds each time. However, for practice data, we only concentrated on the high-risk population in terms of mobility. We purposively

collected information on practices toward COVID-19 only from those respondents who left their homes during the COVID-19 lockdown period. The complete set of data collected is available in [Multimedia Appendix 2](#).

Statistical Analysis

Both descriptive and inferential statistical analyses were performed. In the descriptive analyses, the characteristics of the study participants were presented in terms of frequency and percentage with a 95% confidence interval. The Wilcoxon rank sum test for 2 categorical variables and Kruskal-Wallis test for more than 2 categorical variables were performed to test the statistical inferences across demographic and socioeconomic information and KAP scores, respectively. The Spearman rank correlation test was performed to assess the correlations between the scores of knowledge and attitudes, knowledge and practices, and attitudes and practices. Multiple logistic regression analyses were used to assess the association of the demographic and socioeconomic characteristics of the respondents with their levels of knowledge, attitudes, and practices separately. Furthermore, the associations among knowledge and attitudes, knowledge and practices, and attitudes and practices were assessed to verify the internal consistency. In this analysis, the degree of knowledge, type of attitude, and level of practice were considered dependent variables, whereas age group, sex, education level, occupation, and other variables were considered independent variables. A P value $<.05$ was considered

statistically significant. We analyzed the data using Stata SE, version 13.0 (StataCorp LLC).

Ethics Approval and Consent of Respondents

The research protocol, questionnaire, consent statement, and procedures of obtaining informed consent were approved by the Ethical Review Committee (2020/01) of the Public Health Foundation, Bangladesh. As the interview was conducted over the telephone, informed verbal consent was obtained from all respondents before the interview. Respondents were informed in Bengali about their rights to voluntary participation and to withdraw from the interview at any point. Respondents were assured of the anonymity of the data they provided.

Results

Demographic and Socioeconomic Characteristics

A total of 492 individuals were interviewed in this study, and the majority of them belonged to a younger age group ($n=256$, 52%, aged ≤ 35 years). Among the 492 respondents, 321 (65.2%) were male, 156 (31.7%) had a bachelor's degree or higher level of education, 304 (61.8%) were from a rural area, 202 (41%) were unemployed and had no income, and 449 (91.3%) had access to available running water. Other characteristics are shown in [Table 1](#). Moreover, [Multimedia Appendix 3](#) presents the demographic and socioeconomic characteristics of the respondents by administrative division.

Table 1. Demographic and socioeconomic characteristics of the respondents (N=492).

Characteristic	Values	
	n (%)	95% CI
Age group (years)		
≤25	84 (17.1)	14-20.7
26-35	172 (35)	30.9-39.3
36-45	83 (16.9)	13.8-20.5
46-55	59 (12)	9.4-15.2
56-65	54 (11)	8.5-14.1
≥66	40 (8.1)	6-10.9
Sex		
Male	321 (65.2)	60.9-69.3
Female	171 (34.8)	30.7-39.1
Education level		
No education	76 (15.5)	12.5-18.9
Primary	87 (17.7)	14.5-21.3
Secondary	103 (20.9)	17.6-24.8
Higher secondary	70 (14.2)	11.4-17.6
Bachelor's degree and above	156 (31.7)	27.7-36
Occupation		
Currently unemployed	202 (41.1)	36.8-45.5
Service holder	134 (27.2)	23.5-31.4
Farmer	28 (5.7)	4-8.1
Businessperson	57 (11.6)	9-14.7
Day laborer	52 (10.6)	8.1-13.6
Other	19 (3.9)	2.5-6
Religion		
Muslim	454 (92.3)	89.6-94.3
Other	38 (7.7)	5.7-10.4
Current residence		
Urban	188 (38.2)	34-42.6
Rural	304 (61.8)	57.4-66
Marital status		
In a marital relationship	331 (67.3)	63-71.3
Not in a marital relationship	161 (32.7)	28.7-37
Family size		
1-3 members	109 (22.2)	18.7-26.1
3-6 members	316 (64.2)	59.9-68.4
>7 members	67 (13.6)	10.9-17
Earning person in the family		
None	6 (1.2)	0.5-2.7
1 person	253 (51.4)	47-55.8
≥2 persons	233 (47.4)	43-51.8
Monthly income (BDT^a)		

Characteristic	Values	
	n (%)	95% CI
No income	202 (41.1)	36.8-45.5
≤10,000	111 (22.6)	19.1-26.5
10,001-20,000	80 (16.3)	13.2-19.8
20,001-3,0000	39 (7.9)	5.8-10.7
30,001-40,000	33 (6.7)	4.8-9.3
>40,000	27 (5.5)	3.8-7.9
Availability of running water at home		
Yes	449 (91.3)	88.4-93.5
No	43 (8.7)	6.5-11.6
Division of residence		
Barishal	30 (6.1)	4.3-8.6
Chattogram	87 (17.7)	14.5-21.3
Dhaka	120 (24.4)	20.8-28.4
Khulna	59 (12)	9.4-15.2
Mymensingh	38 (7.7)	5.7-10.4
Rajshahi	69 (14.0)	11.2-17.4
Rangpur	59 (12)	9.4-15.2
Sylhet	30 (6.1)	4.3-8.6

^aBDT: Bangladeshi Taka. 1 BDT=US \$.012.

Knowledge About COVID-19

The mean score of knowledge was estimated at 10.56 (SD 2.86), where the range of scores varied between 0 and 14. The correct response rates regarding the individual knowledge questions ranged between 55.3% (272/492) and 91.5% (450/492). Most of the 492 respondents provided correct responses to the items related to common symptoms (n=447, 90.9%), mode of spreading (n=404, 82.1%), and knowledge about preventive practices, eg, avoiding crowded places (n=409, 83.1%),

maintaining recommended social distancing (n=420, 85.4%), and frequent hand washing (n=450, 91.5%) (a detailed distribution is available in [Multimedia Appendix 4](#)). Overall, 219 of 492 participants (44.5%) were able to provide correct answers for 11 or more questions out of the total 14 questions, for which they obtained a total score of 11.1 (representing a minimum of 80% of the total score). A score of 11.1 or more was considered to indicate good knowledge about COVID-19 ([Table 2](#)).

Table 2. Data related to the survey questions, scores, and levels of knowledge, attitudes, and practices regarding COVID-19 (N=492).

Variable	Number of questions	Score		Level, n (%)	
		Range	Total, mean (SD)	Poor or negative	Good or positive
Knowledge	14	0-14	10.56 (2.86)	273 (55.5)	219 (44.5)
Attitudes	2	0-2	1.24 (0.83)	251 (51)	241 (49)
Practices	5	0-5	3.17 (1.50)	192 (53.5)	167 (46.5)

Attitudes Toward COVID-19

The mean score of attitudes based on the 2 questions in the attitude domain was estimated at 1.24 (SD 0.83, range 0-2). The minimum percentage of respondents providing a positive response for any single question in the attitude domain was 55.3% (272/492), and the maximum percentage of respondents providing a positive response for any single question was 68.5% (337/492). A total of 241 (49%) of the 492 respondents had a strong positive attitude toward the successful control of the pandemic, considering the 2 questions added in the domain.

The maximum possible score was 2, which was 0.76 higher than the mean score in the attitude domain ([Table 2](#)).

Practices Around COVID-19

The mean score of practices was estimated at 3.17 (SD 1.50, range 0-5). The minimum percentage of respondents who provided positive answers for any single question in the practices domain was 42.6% (153/492), and the maximum percentage of respondents providing positive answers for any single question was 76% (273/492) (a detailed distribution is available in [Multimedia Appendix 2](#)). Respondents scoring between 80%

and 100% (equivalent to a score of 4 out of 5) were considered to have good practices toward COVID-19. According to this cutoff, 167 of the 359 respondents (46.5%) who left their homes during the lockdown were found to demonstrate good practices toward COVID-19 (Table 2).

Table 3 shows a comparison of the KAP scores across the respondents' characteristics. A statistically significant knowledge difference was found between age groups, sexes, education levels, occupations, types of residence, marital relationships, and income quantiles, as well as across divisions. Regarding attitudes, statistically significant variations in attitude scores were found between the religious groups, marital relationships, and administrative divisions. Regarding practices, statistical differences were found between age groups, sexes, levels of education, occupations, religions, types of residence,

income quantiles, and water availability at home, as well as across administrative divisions. Similar trends were observed in knowledge and practices: both were comparatively low among the older groups; both knowledge (mean 12.2, SD 1.4; $P < .001$) and practices (mean 5, SD 0.9; $P < .001$) were strongly and positively correlated with a higher level of education. Among different types of professionals, those in the service sector were found to have higher knowledge (mean 12, SD 1.6) and to adopt better practices. On average, rural respondents had significantly lower knowledge (mean 9.8, SD 3.1; $P < .001$) and practice (mean 4, SD 1.4; $P < .001$) scores compared to urban respondents. Regarding the administrative divisions, our results demonstrate that knowledge (mean 7.7, SD 3.6; $P < .001$), attitude (mean 0.8, SD 0.8; $P < .001$), and practice (mean 3.6, SD 1.2; $P < .001$) scores were comparatively low in the Rajshahi division.

Table 3. Mean scores of knowledge, attitude, and practice toward COVID-19 among the respondents.

Characteristic	n	Knowledge score, mean (SD)	P value	Attitude score, mean (SD)	P value	n	Practice score, mean (SD)	P value
Age group^a (years)			<.001		.49			<.001
≤25	84	11.3 (2)		1.3 (0.8)		58	4.1 (1.1)	
26-35	172	11.7 (2.1)		1.3 (0.8)		134	4.7 (1.2)	
36-45	83	10.5 (2.6)		1.3 (0.8)		69	4.2 (1.3)	
46-55	59	9.2 (3)		1.1 (0.9)		46	4 (1.5)	
56-65	54	9.1 (3.4)		1.1 (0.8)		34	4.1 (1.3)	
>66	40	8.1 (3.6)		1.3 (0.8)		18	3.4 (1.6)	
Sex^b			.02		.39			.02
Male	321	10.8 (2.8)		1.3 (0.8)		266	4.5 (1.2)	
Female	171	10.2 (3)		1.2 (0.8)		93	3.9 (1.4)	
Education level^a			<.001		.65			<.001
No education	76	8.4 (3.3)		1.3 (0.8)		45	3.4 (1.2)	
Primary	87	9 (3.5)		1.1 (0.9)		68	3.9 (1.4)	
Secondary	103	10.3 (2.3)		1.2 (0.8)		75	4.1 (1.4)	
Higher secondary	70	11.5 (1.8)		1.3 (0.9)		55	4.5 (1.1)	
Bachelor's degree and above	156	12.2 (1.4)		1.3 (0.8)		116	5 (0.9)	
Occupation^a			<.001		.38			<.001
Currently unemployed	202	10.3 (2.9)		1.2 (0.8)		115	3.8 (1.4)	
Service holder	134	12 (1.6)		1.3 (0.8)		110	5 (0.7)	
Farmer	28	7.6 (3.9)		1 (0.9)		24	4 (1.5)	
Businessperson	57	10.6 (2.4)		1.4 (0.8)		55	4.6 (1.2)	
Day laborer	52	9.4 (2.9)		1.2 (0.9)		39	3.7 (1.4)	
Other	19	10.7 (3.5)		1 (0.9)		16	4.3 (1.2)	
Religion^b			.08		<.001			.19
Muslim	454	10.5 (2.9)		1.2 (0.8)		332	4.4 (1.3)	
Other	38	11.3 (2.3)		1.7 (0.6)		27	3.6 (1.2)	
Current residence^b			<.001		0.28			<.001
Urban	188	11.7 (1.9)		1.3 (0.8)		131	4.8 (0.8)	
Rural	304	9.8 (3.1)		1.2 (0.8)		228	4 (1.4)	
Marital status^b			<.001		.02			.03
In a marital relationship	331	10.1 (2.9)		1.2 (0.8)		250	4.2 (1.4)	
Not in a marital relationship	161	11.5 (2.5)		1.4 (0.8)		109	4.6 (1.1)	
Family size^a			.48		.57			.25
1-3 members	109	10.8 (2.8)		1.2 (0.8)		83	4.4 (1.3)	
3-6 members	316	10.6 (2.9)		1.3 (0.8)		233	4.3 (1.3)	
≥7 members	67	10.2 (2.9)		1.3 (0.8)		43	4 (1.5)	
Earning person in household^a			.30		.24			.48
No earning person	6	12.2 (1.5)		1 (0.9)		3	4.7 (0.6)	
1 earning person	253	10.5 (2.8)		1.2 (0.8)		179	4.3 (1.3)	
≥2 earning persons	233	10.6 (3)		1.3 (0.8)		177	4.3 (1.3)	

Characteristic	n	Knowledge score, mean (SD)	P value	Attitude score, mean (SD)	P value	n	Practice score, mean (SD)	P value
Monthly income^a (BDT^c)			<.001		.13			<.001
No income	202	10.3 (2.9)		1.3 (0.8)		116	3.9 (1.3)	
≤ 10000	111	9.2 (3.3)		1 (0.9)		89	4 (1.4)	
10001-20000	80	11.2 (2.1)		1.4 (0.8)		70	4.8 (1.1)	
20001-30000	39	12.2 (1.9)		1.3 (0.8)		34	4.6 (1.1)	
30001-40000	33	12 (1.5)		1.2 (0.9)		25	5 (0.8)	
> 40000	27	12.5 (2)		1.4 (0.7)		25	4.6 (1.1)	
Availability of water at home^b			.11		.02			.09
Yes	449	10.6 (2.9)		1.3 (0.8)		332	4.3 (1.3)	
No	43	10.4 (2.1)		1 (0.9)		27	4.9 (1)	
Division of residence^a			<.001		<.001			<.001
Barishal	30	10.4 (3)		1.2 (0.8)		19	4.4 (1.5)	
Chattogram	87	11.7 (1.7)		1 (0.9)		62	5 (1.1)	
Dhaka	120	10.8 (2.2)		1.3 (0.7)		85	4.2 (1.2)	
Khulna	59	11.8 (1.3)		1.8 (0.5)		53	4.8 (1)	
Mymensingh	38	11 (2.6)		1.6 (0.8)		29	4.6 (1.7)	
Rajshahi	69	7.7 (3.6)		0.8 (0.8)		68	3.6 (1.2)	
Rangpur	59	9.4 (3.6)		1.3 (0.8)		26	3.8 (1.4)	
Sylhet	30	12.3 (1.6)		1.5 (0.9)		17	3.9 (1.3)	

^aKruskal-Wallis test.

^bWilcoxon rank sum test.

^cBDT: Bangladeshi Taka. 1 BDT=US \$.012.

Correlations Between Knowledge, Attitude, and Practice Scores

Table 4 illustrates the Spearman rank correlation coefficients and the statistically significant levels of the coefficients. The

Table 4. Correlations between knowledge, attitude, and practice scores.

Correlation	Spearman rank correlation coefficient	P value
Knowledge-Attitude	0.21	<.001
Knowledge-Practice	0.45	<.001
Attitude-Practice	0.27	<.001

Factors Associated With Knowledge

Model 1 in Table 5 demonstrates the factors associated with knowledge about COVID-19. The unadjusted model showed that several factors, such as age group, sex, education, occupation, current residence, administrative division, marital status, and income level were significantly associated with knowledge on COVID-19; the adjusted model, after removing the confounding effects of the independent variables, showed that education, marital status, family size, monthly income, and administrative divisions were significantly associated with knowledge about COVID-19. The higher age groups (ie, 46-55 years: odds ratio [OR] 0.34, 95% CI 0.17-0.7; 56-65 years: OR

test showed positive and statistically strong correlations of knowledge with attitudes, knowledge with practices, and attitudes with practices ($P<.001$).

0.42, 95% CI 0.2-0.87; ≥66 years: OR 0.18, 95% CI 0.07-0.46) were more likely to have poor knowledge compared to the reference age group of 25 years or below. Female respondents (OR 0.64; 95% CI 0.44-0.94) were more likely to have poor knowledge compared to their male counterparts. Among the different occupation groups, farmers (OR 0.42, 95% CI 0.16-1.09) and day laborers (OR 0.42, 95% CI 0.2-0.86) were significantly more likely to have lower knowledge compared to those who were not employed. In terms of residence, those living in rural areas had a significantly lower level of knowledge toward COVID-19 (OR 0.44, 95% CI 0.3-0.64) compared to their urban counterparts. According to administrative divisions, the Rajshahi division respondents (OR 0.25, 95% CI 0.09-0.65,

and adjusted OR 0.22, 95% CI 0.06-0.73) had a lower level of knowledge compared to the respondents from the reference division, Barishal.

Table 5. Association of characteristics with knowledge and attitudes toward COVID-19.

Characteristic	Model 1: Knowledge		Model 2: Attitudes	
	OR ^a (95% CI)	Adjusted OR (95% CI)	OR (95% CI)	Adjusted OR (95% CI)
Age group (years)				
≤25	Ref ^b	Ref	Ref	Ref
26-35	1.77* (1.05-3.01)	1.93 (0.92-4.06)	0.98 (0.58-1.64)	0.95 (0.48-1.89)
36-45	0.57 (0.3-1.05)	0.79 (0.31-1.99)	0.98 (0.53-1.79)	1.15 (0.5-2.61)
46-55	0.34* (0.17-0.7)	0.87 (0.31-2.39)	0.75 (0.38-1.47)	0.74 (0.3-1.8)
56-65	0.42* (0.2-0.87)	1.33 (0.45-3.9)	0.66 (0.33-1.31)	0.66 (0.25-1.72)
≥66	0.18* (0.07-0.46)	0.38 (0.1-1.41)	1.05 (0.5-2.24)	0.79 (0.28-2.26)
Sex				
Male	Ref	Ref	Ref	Ref
Female	0.64* (0.44-0.94)	0.67 (0.38-1.2)	0.87 (0.6-1.27)	0.9 (0.54-1.5)
Education level				
No education	Ref	Ref	Ref	Ref
Primary	2.38* (1.08-5.25)	1.88 (0.75-4.69)	0.77 (0.42-1.43)	0.74 (0.35-1.56)
Secondary	2.43* (1.13-5.23)	1.59 (0.62-4.06)	0.8 (0.44-1.44)	0.56 (0.26-1.2)
Higher secondary	7.88* (3.56-17.45)	4.58* (1.53-13.72)	1.19 (0.62-2.29)	0.57 (0.22-1.5)
Bachelor's degree and above	15.53* (7.49-32.2)	6.07* (2.14-17.23)	0.92 (0.53-1.6)	0.3* (0.12-0.77)
Occupation				
Currently unemployed	Ref	Ref	Ref	Ref
Service holder	2.98* (1.89-4.69)	0.77 (0.27-2.2)	1.16 (0.75-1.8)	2.2 (0.83-5.81)
Farmer	0.42 (0.16-1.09)	0.62 (0.16-2.45)	0.73 (0.33-1.63)	1.58 (0.46-5.42)
Businessperson	1.22 (0.67-2.2)	0.43 (0.13-1.39)	1.66 (0.92-3.02)	3.54* (1.2-10.47)
Day laborer	0.42* (0.2-0.86)	0.35 (0.11-1.14)	1.04 (0.57-1.92)	2.37 (0.85-6.61)
Other	1.73 (0.67-4.45)	1.15 (0.29-4.6)	0.82 (0.32-2.12)	1.85 (0.48-7.19)
Religion				
Muslim	Ref	Ref	Ref	Ref
Other	1.6 (0.82-3.11)	0.8 (0.29-2.2)	5.15* (2.22-11.93)	3.56* (1.31-9.64)
Current residence				
Urban	Ref.	Ref.	Ref.	Ref.
Rural	0.44* (0.3-0.64)	0.67 (0.39-1.14)	0.87 (0.61-1.26)	0.77 (0.47-1.27)
Marital status				
In a marital relationship	Ref	Ref	Ref	Ref
Not in a marital relationship	2.6* (1.76-3.82)	1.98* (1.14-3.43)	1.4 (0.96-2.05)	1.56 (0.95-2.57)
Family size				
1-3 members	Ref	Ref	Ref	Ref
3-6 members	0.98 (0.63-1.51)	2.13* (1.14-3.96)	1.32 (0.85-2.05)	1.49 (0.86-2.61)
>7 members	0.56 (0.3-1.05)	0.98 (0.39-2.47)	1.16 (0.63-2.14)	1.27 (0.57-2.83)
Earning person in household				
None	Ref	Ref	Ref	Ref
1	0.38 (0.07-2.1)	0.23 (0.03-1.81)	1.72 (0.31-9.56)	2.95 (0.46-18.95)
≥2	0.42 (0.07-2.32)	0.23 (0.03-1.82)	2.2 (0.39-12.24)	4.03 (0.61-26.56)

Characteristic	Model 1: Knowledge		Model 2: Attitudes	
	OR ^a (95% CI)	Adjusted OR (95% CI)	OR (95% CI)	Adjusted OR (95% CI)
Monthly income (BDT^c)				
No income	Ref	Ref	Ref	Ref
≤ 10000	0.81 (0.5-1.33)	2.2 (0.82-5.92)	0.74 (0.46-1.18)	0.42 (0.17-1.03)
10001-20000	1.6 (0.95-2.7)	1.56 (0.54-4.5)	1.34 (0.79-2.25)	0.37 (0.14-1.01)
20001-30000	4.5* (2.12-9.56)	2.51 (0.74-8.54)	1.1 (0.55-2.17)	0.29* (0.09-0.91)
30001-40000	3.53* (1.62-7.7)	2.73 (0.78-9.58)	1.11 (0.53-2.31)	0.32 (0.1-1.06)
>40000	10.16* (3.38-30.52)	14.28* (3.08-66.16)	1.12 (0.5-2.5)	0.25* (0.07-0.9)
Availability of water at home				
Yes	Ref	Ref	Ref	Ref
No	0.51 (0.26-1.01)	0.54 (0.24-1.24)	0.53 (0.27-1.02)	0.54 (0.25-1.15)
Knowledge about COVID-19				
Poor	N/A ^d	N/A	Ref	Ref
Good	N/A	N/A	1.93* (1.34-2.76)	1.76* (1.07-2.89)
Division of residence				
Barishal	Ref	Ref	Ref	Ref
Chattogram	1.4 (0.61-3.23)	1.38 (0.48-3.94)	1.06 (0.45-2.47)	1.14 (0.46-2.84)
Dhaka	0.7 (0.31-1.59)	1.02 (0.36-2.91)	1.11 (0.49-2.51)	0.86 (0.35-2.12)
Khulna	2.37 (0.96-5.81)	1.58 (0.51-4.89)	7.35* (2.71-19.94)	8.96* (2.98-26.98)
Mymensingh	2.01 (0.76-5.3)	2.87 (0.83-9.94)	4.2* (1.5-11.73)	4.29* (1.39-13.27)
Rajshahi	0.25* (0.09-0.65)	0.22* (0.06-0.73)	0.45 (0.18-1.14)	0.42 (0.15-1.18)
Rangpur	0.84 (0.34-2.04)	3.41 (0.98-11.83)	1.35 (0.56-3.3)	1.54 (0.53-4.49)
Sylhet	5.23* (1.66-16.51)	5.37* (1.24-23.22)	3.5* (1.2-10.2)	1.63 (0.49-5.42)

^aOR: odds ratio.

^bRef: reference.

^cBDT: Bangladeshi Taka. 1 BDT=US \$.012.

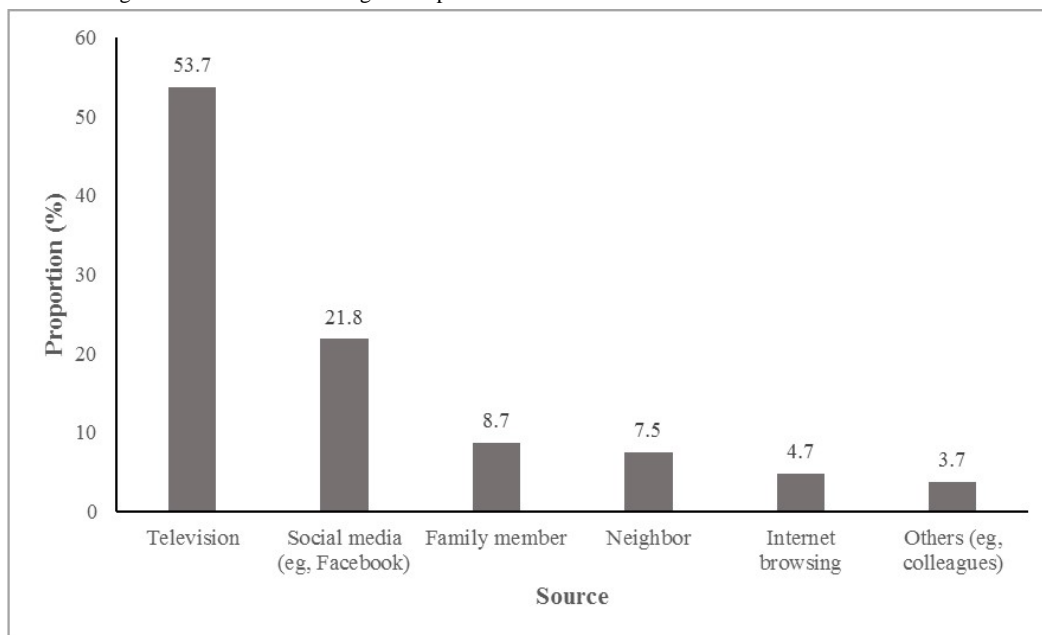
^dN/A: not applicable.

**P*<.05.

Figure 2 shows the sources of knowledge and indicates that television (264/492, 53.7%), followed by social media (107/492; 21.8%), is the major source of knowledge on COVID-19 for

the study population. Other important sources included family members (43/492, 8.7%), neighbors (37/492, 7.5%), and the internet (23/492, 4.7%).

Figure 2. Source of knowledge about COVID-19 among the respondents.



Factors Associated With Attitudes

Model 2 of Table 5 demonstrates the factors associated with attitudes toward COVID-19. In the unadjusted model, religion, knowledge about COVID-19, and administrative division showed a significant association with a positive attitude toward COVID-19. Controlling for the confounding effects of the independent variables confirmed the significant influence of education, occupation, religion, monthly income, knowledge about COVID-19, and administrative division on positive attitudes toward COVID-19. According to education level, people with higher education levels (adjusted OR 0.3, 95% CI 0.12-0.77) were more likely to have a negative attitude toward

control of COVID-19 than those with no education. People with an income level between BDT 20,001 and 30,000 (US \$198 and \$297, adjusted OR 0.29, 95% CI 0.09-0.91) and more than BDT 40,000 (US \$396, adjusted OR 0.25, 95% CI 0.07-0.9) had a higher likelihood of a negative attitudes toward controlling COVID-19 compared to those with no income.

Figure 3 shows that 359 of the 492 study participants (73%) went outside the home in the week preceding the survey during the lockdown period. As mentioned earlier in the Methods section, information on practices toward COVID-19 was collected only for those respondents who left their homes during the lockdown period. Among this group, 266 of 359 (74.1%) were male.

Figure 3. Distribution of respondents according to their mobility during the lockdown period.

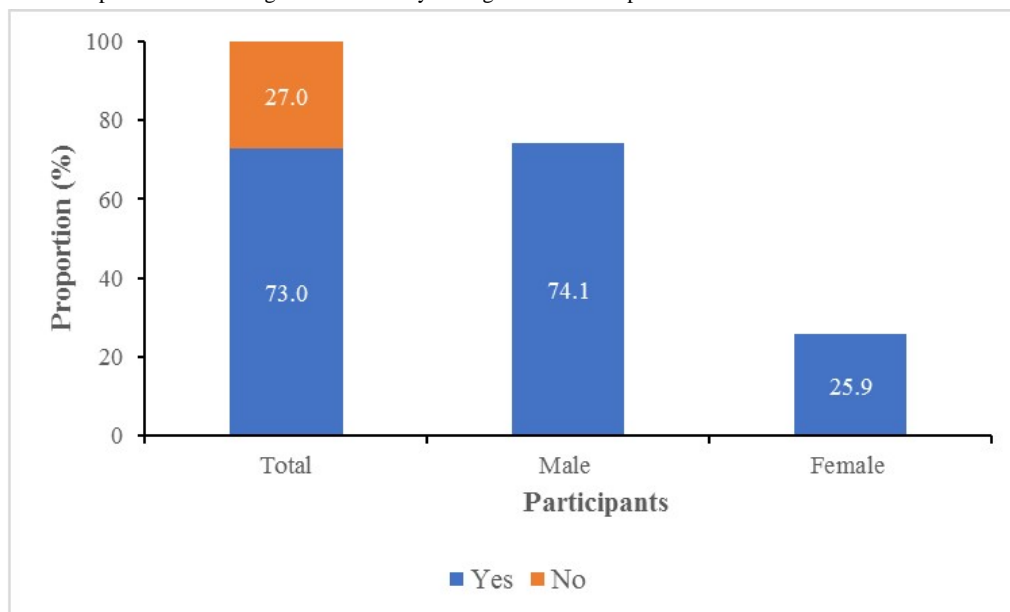
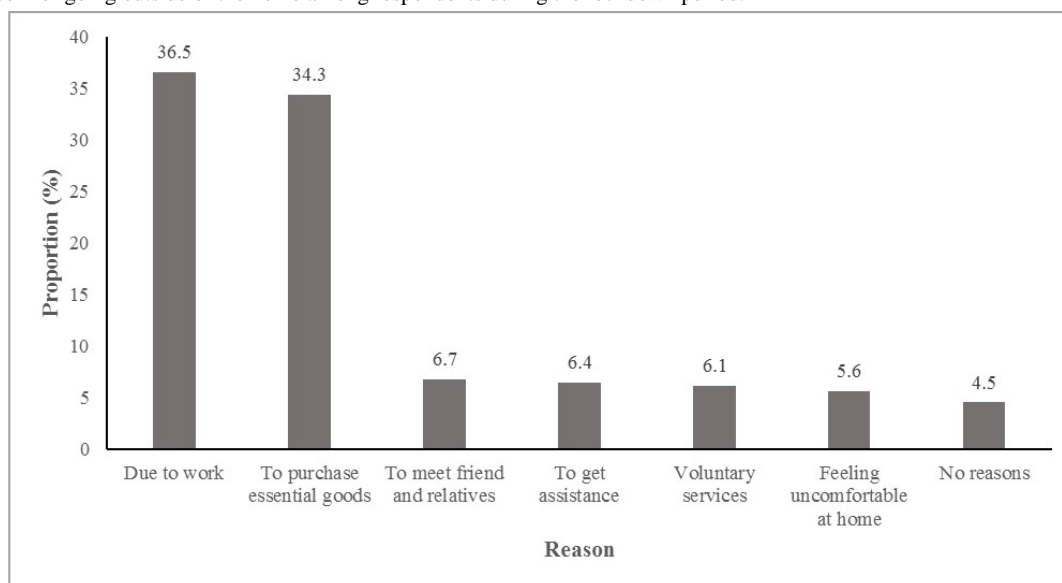


Figure 4 shows that among the 359 respondents who left their homes during the lockdown, 131 (37.5%) went out due to work, followed by 123 (34.3%) to purchase essential goods such as

food and medicine. However, 16 of the 359 respondents (4.5%) went out for no particular reason.

Figure 4. Reason for going outside of the home among respondents during the lockdown period.

Factors Associated With Practices

Table 6 demonstrates the factors associated with self-reported practices regarding COVID-19. The unadjusted model showed that several sociodemographic factors, such as age group, education, occupation, monthly income, water availability at home, knowledge, attitude, and administrative division were significantly associated with good practices toward COVID-19. The adjusted model, after removing the other variables' confounding effects, showed that education level, occupation, religion, monthly income, knowledge, attitude, and

administrative divisions were significantly associated with good practices toward COVID-19. The results showed that compared to Muslims, people of other religions (adjusted OR 0.14; 95% CI 0.04-0.51) had a higher likelihood of poor practices. According to income level, respondents with a monthly income of more than BDT 40,000 (US \$396; adjusted OR 0.1; 95% CI 0.02-0.6) had more likelihood of poor practices than those in the no income reference category. Respondents who were from the Rajshahi division (OR 0.18; 95% CI 0.06-0.59) were more likely to have poor practices compared to those from the reference Barisal division.

Table 6. Association of characteristics with self-reported practices toward COVID-19.

Characteristic	n	Unadjusted OR ^a (95% CI)	Adjusted OR (95% CI)
Age group (years)			
≤25	58	Ref ^b	Ref
26-35	134	2.06* (1.1-3.85)	1.52 (0.56-4.12)
36-45	69	1.57 (0.77-3.18)	1.21 (0.36-4)
46-55	46	0.98 (0.44-2.16)	1.93 (0.5-7.41)
56-65	34	0.73 (0.3-1.77)	2.27 (0.55-9.4)
≥66	18	0.3 (0.08-1.17)	0.52 (0.08-3.19)
Sex			
Male	266	Ref	Ref
Female	93	0.69 (0.43-1.12)	0.92 (0.39-2.18)
Education level			
No education	45	Ref	Ref
Primary	68	1.46 (0.61-3.5)	1.03 (0.33-3.22)
Secondary	75	2.9* (1.26-6.7)	2.75 (0.86-8.82)
Higher secondary	55	2.92* (1.21-7.04)	1.89 (0.49-7.22)
Bachelor's degree and above	116	7.18* (3.22-16.03)	7.43* (1.95-28.33)
Occupation			
Currently unemployed	115	Ref	Ref
Service holder	110	2.51* (1.47-4.3)	1.99 (0.42-9.47)
Farmer	24	0.97 (0.39-2.4)	7.34* (1.09-49.33)
Businessperson	55	1.67 (0.87-3.2)	2.39 (0.45-12.75)
Day laborer	39	0.72 (0.33-1.56)	3.47 (0.67-18.09)
Other	16	1.26 (0.44-3.61)	1.3 (0.22-7.86)
Religion			
Muslim	332	Ref	Ref
Other	27	0.46 (0.2-1.08)	0.14* (0.04-0.51)
Current residence			
Urban	131	Ref	Ref
Rural	228	0.71 (0.46-1.09)	0.94 (0.46-1.93)
Marital status			
In a marital relationship	250	Ref	Ref
Not in a marital relationship	109	1.07 (0.68-1.68)	0.89 (0.43-1.85)
Family size			
1-3 members	83	Ref	Ref
3-6 members	233	1.04 (0.63-1.72)	0.82 (0.38-1.8)
≥7 members	43	0.67 (0.31-1.42)	0.68 (0.21-2.25)
Earning person in household			
None	3	Ref.	Ref.
1 person	179	0.51 (0.05-5.68)	0.21 (0.01-3.6)
≥2 persons	177	0.37 (0.03-4.13)	0.19 (0.01-3.15)
Monthly income (BDT^c)			

Characteristic	n	Unadjusted OR ^a (95% CI)	Adjusted OR (95% CI)
No income	116	Ref	Ref
≤10000	89	0.65 (0.36-1.16)	0.32 (0.08-1.33)
10001-20000	70	2.4* (1.3-4.41)	0.9 (0.21-3.8)
20001-30000	34	1.79 (0.83-3.88)	0.32 (0.06-1.74)
30001-40000	25	3.0*1 (1.2-7.54)	0.85 (0.13-5.77)
>40000	25	1.11 (0.47-2.66)	0.1* (0.02-0.6)
Availability of water at home			
Yes	332	Ref	Ref
No	27	2.46* (1.07-5.63)	2.27 (0.68-7.6)
Knowledge about COVID-19			
Poor	195	Ref	Ref
Good	164	3.13* (2.03-4.83)	2.33* (1.16-4.68)
Attitude toward COVID-19			
Negative	96	Ref	Ref
Positive	263	2.43* (1.59-3.72)	3.87* (1.95-7.68)
Division of residence			
Barishal	19	Ref	Ref
Chattogram	62	6.38* (2.08-19.53)	11.85* (2.72-51.57)
Dhaka	85	0.87 (0.32-2.4)	1.5 (0.4-5.63)
Khulna	53	6.72* (2.11-21.42)	10.09* (2.15-47.41)
Mymensingh	29	1.47 (0.46-4.73)	0.89 (0.2-4.03)
Rajshahi	68	0.18* (0.06-0.59)	0.32 (0.07-1.42)
Rangpur	26	0.33 (0.09-1.24)	0.48 (0.09-2.74)
Sylhet	17	0.29 (0.06-1.38)	0.24 (0.04-1.51)

^aOR: odds ratio.

^bRef: reference.

^cBDT: Bangladeshi Taka. 1 BDT=US \$.012.

**P*<.05.

Discussion

Principal Results

Appropriate participation is essential for effective control and prevention of an infectious disease such as COVID-19. Nevertheless, this participation relies fundamentally on adequate knowledge and improved practices against the infection. In our study, we applied a survey through one-to-one audio communication instead of web-based surveys to evaluate gaps and needs in the fight against COVID-19 during the period of its rapid rise in Bangladesh. We found that 55.5% of respondents had poor knowledge, 49% expressed a negative attitude about controlling this disease nationally and globally, and 53.5% had unfavorable practices toward COVID-19. Our analysis has shown that the knowledge related to COVID-19 of certain demographic and socioeconomic groups (eg, those who were aged ≥46 years; were female; had no education; were farmers, day laborers, or rural residents; were in a marital relationship; had a larger family; earned less than BDT 20,001 [US \$236];

and were residents of Rajshahi Division) is significantly lower than that of respondents in the corresponding reference categories. Most people rely on television, followed by social media, as a source of knowledge. Almost three-quarters of the respondents left home during the lockdown period, the majority of whom were male (74%). Those with no education lagged behind those with secondary or higher education. Finally, we found that better practice of COVID-19 health measures was associated with good knowledge and a positive attitude toward COVID-19.

We also found poor knowledge to be significantly associated with several demographic and socioeconomic factors of individuals. Better socioeconomic status contributed to a higher level of COVID-19-related knowledge among people in Bangladesh. Although the study was conducted some months after the onset of the pandemic, the influence of socioeconomic status remained, indicating unequal penetration of information on COVID-19 management throughout society. The study findings indicate that knowledge about COVID-19 increased

as age decreased; older people had lower levels of knowledge. This finding is supported by several other international studies from developing and developed countries that report older respondents having poor knowledge of COVID-19 compared to younger populations [10,21,25]. The physical condition, cognitive status, and limited access to information of the older population may influence their ability to understand information on COVID-19 and may contribute to the poorer knowledge level among this population [26]. Lack of familiarity and limited use of modern technology may be other reasons for poor knowledge among older adults [27,28]. The study found that farmers and daily laborers were more likely to have poor knowledge about COVID-19. This finding is similar to that in a study conducted in China [21], which found that laborers had poorer knowledge of this topic.

From the beginning of the pandemic, there has never been an effective lockdown in place in Bangladesh. Primarily, this may have been the result of the government's policy to declare this crucial time a general holiday rather than a lockdown [29,30]. The notion of a holiday did not help people to recognize the importance of staying at home as much as possible. People were under the impression that they were free to roam around. As a result, many people willingly ignored the stay-at-home or social distancing guidelines. They took the opportunity to move to different cities across the country, which we anticipate might have contributed to the rapid spread of infection at the community level throughout the country [31]. Subsequently, the government extended the general holiday without ensuring adequate subsistence support for low-income residents before imposing a lockdown, which may also have compelled people to leave their homes [32]. Further, the rapid change in lockdown timing every week might have precluded people from making adequate preparation. Moreover, the government's inability to provide information on how people in lockdown situations could access essential materials for their everyday life and engage community groups for meeting essential needs may be one reason for the poor practice of safety measures among the population [33]. This result reinforces the conclusions of previous studies identifying strict prevention practices and mobilizing community volunteers to support people under lockdown as the primary solutions for reducing the spread of COVID-19 in China and Vietnam [21,34,35].

Bangladesh is still a predominantly rural country, with only 37% of its population living in urban areas [36]. However, rural areas lag behind urban areas in most socioeconomic and health indicators. For example, only 76% of rural areas are served by the national electricity grid (compared to 92% of urban areas), and only 38% of rural households own a television (compared to 70% in urban areas). In terms of demographic and health indicators, we find that 61% of rural women marry before the age of 18 years (compared to 55% of urban women), 79% of rural women receive antenatal care from a skilled provider (compared to 90% of urban women), and 45% give birth in a health facility (compared to 63% of urban women); moreover, 33% of rural children under the age of 5 years are stunted (compared to 25% of urban children) [37]. Consistent with this reduced health-seeking behavior, the current study found that rural residents were less likely to adhere to COVID-19-related

hygiene practices. Adding to this poor practice was the risk of the rapid spread of infection caused by migrants returning to the rural community from different countries around the world during the pandemic period [38,39]. Lower awareness in rural communities about health policy and programs has been observed in several other countries, and this phenomenon may pose a higher degree of threat in the case of communicable diseases such as COVID-19 [10,40,41].

Good knowledge and a positive attitude toward controlling COVID-19 were found to be associated with good practice of safety measures. This finding is well recognized in several global studies, in which it was found that good knowledge and a positive attitude toward COVID-19 lead to improved practice of safety measures [10,22,42,43]. It is worth mentioning that the consistency of theory-based approaches demonstrates an association among knowledge, belief, and change in human behavior [44]. Thus, it has been recognized that adequate and proper knowledge of a specific health emergency is a key modifier of personal belief that leads to changing human behavior [45]. When combined with good knowledge and positive attitudes, a higher level of appropriate practice was identified among respondents. This may result from a higher level of accurate knowledge about common symptoms, mode of spread, and knowledge about preventive practices (eg, avoiding crowded places, maintaining recommended social distance, and frequent hand washing).

Because KAP levels vary across different socioeconomic groups, we recommend that customized information on COVID-19 be developed targeting different groups in Bangladesh, such as villagers, slum-dwellers, township residents, and urban middle-class citizens. Special emphasis should be given to the groups with lower KAP scores, such as older persons, females, less educated people, farmers, day laborers, rural residents, those in a marital relationship, those with a larger family, those with low income, and residents of the Rajshahi division. The information should be clearly and widely circulated through contextually appropriate channels at a massive level, emphasizing television and social media, as these were found to be the major sources of information. Additionally, because many people did not comply with the lockdown directives, the implementation of a lockdown should encompass a well-planned support system that includes subsistence support for low-income individuals, arranging emergency requirements for the locked-down community, communicating clearly what to do and what not to do during the lockdown, and clarifying whom to consult in the case of any unforeseen situation. A voluntary community support group might be engaged to support people's needs. Instead of arbitrarily increasing the duration of lockdown on a week-by-week basis, a predetermined guide period should be established, in consultation with epidemiologists, and clearly communicated to the public so that they may make adequate preparation to stay at home for the entire period. The term "national holiday" may not convey the right message to the public; therefore, "lockdown" or any contextually appropriate synonym, in consultation with communication experts or social scientists, should be used instead. Special attention should be directed toward rural communities, where engagement in COVID-19 health practices is lower. Finally, because practices

are associated with knowledge and attitude, we recommend that a scientifically oriented social and behavioral change communication strategy be developed in consultation with the relevant experts. Religious, cultural, political, and other community-based agents can be consulted and actively engaged in turning these strategies into actions or practices.

Limitations

This study is not free from limitations. First, the study is limited by the sampling method. Thus, selection bias may exist. Another limitation is the study population. This study only considered adults who use mobile phones. Moreover, data were collected over mobile phone calls, which is not a familiar process in Bangladesh. Thus, response bias may exist, and only those concerned more about the COVID-19 emergency may have participated in the study. Therefore, the results may not be generalizable to other people who are not mobile phone users. The number of questions under the Attitudes section, where only 2 questions were considered in the KAP questionnaire to measure the attitude level, limited the depth of this domain. The major limitation can be considered with regard to the study design. As this is a cross-sectional study, causal inferences cannot be drawn between the significant factors and the KAP levels.

Strengths

Despite the above limitations, the study's strength lies in the fact that it includes data from 8 administrative divisions

throughout the country and that participants from both rural and urban areas were surveyed by telephone. The findings of the study, to our understanding, will motivate and alert policy makers and program implementers who are working on appropriate health education, risk communication, community engagement, and social and behavioral change communication strategies based on levels of KAP toward COVID-19.

Further research is needed to understand the KAP of service providers in COVID-19 pandemic response. Qualitative formative research will aid the design of effective communication strategies to address the pandemic, and subsequent implementation and evaluation research can generate useful knowledge about the implementation and scaleup of such strategies in different parts of Bangladesh and other countries facing similar KAP-related challenges.

Conclusions

Risk communication and community engagement are an integral part of pandemic management [46]. For countries such as Bangladesh, where most of the community lacks proper health practices and where timely information penetration is unequal, KAP studies such as this one can support public health decision-makers in designing evidence-informed and context-specific risk communication and behavioral change communication strategies. These evidence-informed strategies may help in managing health emergencies such as the COVID-19 pandemic for people from all socioeconomic groups.

Acknowledgments

The authors are thankful to the Public Health Foundation, Bangladesh, and its members for providing scientific guidance for conducting the research. The authors are also thankful to Dr Dipak Kumar Mitra, Professor and Chairman, Department of Public Health, School of Health and Life Sciences, North South University, and Professor Sharmeen Yasmeen, Head of the Department, Department of Community Medicine, Bangladesh Medical College, for their valuable inputs in this study. Mr Quazi Maksudur Rahman, Mr Sourav Paul, Md Abdur Razzak, Md Maksudul Alam, Md Asmaul Hasan Khan, Md Jahidul Islam Khan, Md Tarekul Islam, Md Nazrul Islam, Md Ayub Ali, Irin Akter, Pritidipto Paul Chowdhury, and Md Sarwar Jahan Bhuiyan contributed to the telephonic interviews. The authors are grateful to Mr Nicholas Freeland, International Freelance Consultant (Social Protection), for his linguistic, grammatical, and punctuational assistance during the revision of this manuscript. Publication of this article was cosponsored by the *Journal of Medical Internet Research*. All authors declared that they failed to identify any funding source for the project described in the paper to enable coverage of the publication costs. All department/division heads/chairs of the author institutions have declared that no institutional funding mechanisms for knowledge dissemination activities and/or coverage of open access publishing costs at their department or university level exist.

Authors' Contributions

MGR, OA, and TJ contributed to conceptualization, methodological design, and investigation. MGR and MZH were responsible for the analysis plan. MGR, OA, MZH, NS, SSM, and TJ contributed to the writing, review, and editing of the original manuscript. TJ supervised the study team. All authors have approved the final manuscript and agreed to its publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study questionnaire.

[[DOCX File , 45 KB - formative_v5i11e28344_app1.docx](#)]

Multimedia Appendix 2

Study data set.

[[XLSX File \(Microsoft Excel File\), 78 KB](#) - [formative_v5i11e28344_app2.xlsx](#)]

Multimedia Appendix 3

Cross-tabulation of demographic and socioeconomic characteristics of respondents by administrative division.

[[DOCX File , 48 KB](#) - [formative_v5i11e28344_app3.docx](#)]

Multimedia Appendix 4

Response of questions regarding knowledge, attitudes, and practices toward COVID-19.

[[DOCX File , 35 KB](#) - [formative_v5i11e28344_app4.docx](#)]

References

1. Novel Coronavirus (2019-nCoV) Situation Report - 1. World Health Organization. 2020 Jan 21. URL: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10> [accessed 2021-10-15]
2. Weekly epidemiological update on COVID-19. World Health Organization. 2021 Mar 23. URL: <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---23-march-2021> [accessed 2021-03-25]
3. COVID-19 Coronavirus Pandemic. Worldometer. URL: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdUOA?Si [accessed 2021-03-25]
4. COVID-19 Morbidity and Mortality Weekly Update No 53. World Health Organization Bangladesh. 2021 Mar 22. URL: https://cdn.who.int/media/docs/default-source/searo/bangladesh/covid-19-who-bangladesh-situation-reports/who_covid-19-update_56_20210322.pdf?sfvrsn=479b87bc_9 [accessed 2021-10-15]
5. COVID-19 Situation Report No. 10. World Health Organization Bangladesh. 2020 Apr 05. URL: https://www.who.int/docs/default-source/searo/bangladesh/covid-19-who-bangladesh-situation-reports/who-ban-covid-19-sitrep-10.pdf?sfvrsn=c0aac0b8_4 [accessed 2021-10-15]
6. Beyond recovery: towards 2030. United Nations Development Programme. URL: <https://www.bd.undp.org/content/bangladesh/en/home/coronavirus/support-to-national-response.html> [accessed 2021-03-27]
7. Report on Bangladesh Sample Vital Statistics. Statistics and Informatics Division, Ministry Of Planning, Bangladesh. 2018. URL: https://bbs.portal.gov.bd/sites/default/files/files/bbs.portal.gov.bd/page/6a40a397_6ef7_48a3_80b3_78b8d1223e3f/SVRS_Report_2018_29-05-2019%28Final%29.pdf [accessed 2021-10-18]
8. Country Profile: Bangladesh - World Development Indicators. World Bank. 2019. URL: <https://data.worldbank.org/country/bangladesh?view=chart> [accessed 2021-10-18]
9. Poverty data: Bangladesh. Asian Development Bank. 2020. URL: <https://www.adb.org/countries/bangladesh/poverty> [accessed 2020-10-27]
10. Akalu Y, Ayelign B, Molla MD. Knowledge, attitude and practice towards COVID-19 among chronic disease patients at Addis Zemen Hospital, Northwest Ethiopia. *Infect Drug Resist* 2020;13:1949-1960 [FREE Full text] [doi: [10.2147/IDR.S258736](https://doi.org/10.2147/IDR.S258736)] [Medline: [32612371](https://pubmed.ncbi.nlm.nih.gov/32612371/)]
11. Lau LL, Hung N, Go DJ, Ferma J, Choi M, Dodd W, et al. Knowledge, attitudes and practices of COVID-19 among income-poor households in the Philippines: a cross-sectional study. *J Glob Health* 2020 Jun;10(1):011007 [FREE Full text] [doi: [10.7189/jogh.10.011007](https://doi.org/10.7189/jogh.10.011007)] [Medline: [32566169](https://pubmed.ncbi.nlm.nih.gov/32566169/)]
12. The DS. Coronavirus outbreak: govt orders closure of public, private offices from March 26 to April 4. *The Daily Star*. 2020 Mar 23. URL: <https://www.thedailystar.net/coronavirus-deadly-new-threat/news/govt-offices-closed-march-26-april-4-cabinet-secretary-1884730> [accessed 2020-03-26]
13. Coronavirus outbreak: shutdown won't be extended after May 30. *The Daily Star*. 2020 May 27. URL: <https://www.thedailystar.net/coronavirus-outbreak-shutdown-wont-be-extended-after-may-30-1905826> [accessed 2020-05-27]
14. Dhaka T. Restriction on public movement extended till August 31. *Dhaka Tribune*. 2020 Aug 03. URL: <https://www.dhakatribune.com/health/coronavirus/2020/08/03/restriction-on-public-movement-extended-till-august-31> [accessed 2020-08-03]
15. Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations. World Health Organization. 2020 Mar 29. URL: https://webcache.googleusercontent.com/search?q=cache:xIe2Mz4qLr4J:https://www.who.int/docs/default-source/coronaviruse/20200329-scientific-brief1-final-corrected2-revised-final.pdf%3Fsfvrsn%3Dcb817b3a_6%26download%3Dtrue+%cd=1&hl=en&ct=clnk&gl=bd [accessed 2021-10-18]
16. Jefferson T, Del Mar CB, Dooley L, Ferroni E, Al-Ansary LA, Bawazeer GA, et al. Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database Syst Rev* 2011 Jul 06(7):CD006207 [FREE Full text] [doi: [10.1002/14651858.CD006207.pub4](https://doi.org/10.1002/14651858.CD006207.pub4)] [Medline: [21735402](https://pubmed.ncbi.nlm.nih.gov/21735402/)]
17. Gamma AE, Slekiene J, Mosler H. The impact of various promotional activities on Ebola prevention behaviors and psychosocial factors predicting Ebola prevention behaviors in the Gambia evaluation of Ebola prevention promotions. *Int J Environ Res Public Health* 2019 Jun 06;16(11):2020 [FREE Full text] [doi: [10.3390/ijerph16112020](https://doi.org/10.3390/ijerph16112020)] [Medline: [31174301](https://pubmed.ncbi.nlm.nih.gov/31174301/)]

18. Melesie Taye G, Bose L, Beressa TB, Tefera GM, Mosisa B, Dinsa H, et al. COVID-19 knowledge, attitudes, and prevention practices among people with hypertension and diabetes mellitus attending public health facilities in Ambo, Ethiopia. *Infect Drug Resist* 2020;13:4203-4214 [FREE Full text] [doi: [10.2147/IDR.S283999](https://doi.org/10.2147/IDR.S283999)] [Medline: [33262615](https://pubmed.ncbi.nlm.nih.gov/33262615/)]
19. Bangladesh Multiple Indicator Cluster Survey 2019, Key Findings. Progotir Pathey, Bangladesh. 2019. URL: http://bbs.portal.gov.bd/sites/default/files/files/bbs.portal.gov.bd/page/b343a8b4_956b_45ca_872f_4cf9b2f1a6e0/37817b8e25d0d6c1f442e294921ff85e.pdf [accessed 2021-10-18]
20. National Guidelines on Clinical Management of Coronavirus Disease (Covid-19). Government of the People's Republic of Bangladesh. 2019. URL: https://covidlawlab.org/wp-content/uploads/2021/01/Bangladesh_2020.11.05_Guideline_National-Guidelines-on-Clinical-Management-of-COVID-19_EN.pdf [accessed 2021-10-18]
21. Zhong B, Luo W, Li H, Zhang Q, Liu X, Li W, et al. Knowledge, attitudes, and practices towards COVID-19 among Chinese residents during the rapid rise period of the COVID-19 outbreak: a quick online cross-sectional survey. *Int J Biol Sci* 2020;16(10):1745-1752 [FREE Full text] [doi: [10.7150/ijbs.45221](https://doi.org/10.7150/ijbs.45221)] [Medline: [32226294](https://pubmed.ncbi.nlm.nih.gov/32226294/)]
22. Erfani A, Shahriarirad R, Ranjbar K, Mirahmadizadeh A, Moghadami M. Knowledge, attitude and practice toward the novel coronavirus (COVID-19) outbreak: a population-based survey in Iran. *Bull World Health Organ Preprint* posted online on March 30, 2020. [doi: [10.2471/BLT.20.256651](https://doi.org/10.2471/BLT.20.256651)]
23. 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)]
24. Bloom BS. Learning for mastery. *Eval Comment* 1968 May;1(2) [FREE Full text]
25. Wolf MS, Serper M, Opsasnick L, O'Connor RM, Curtis LM, Benavente JY, et al. Awareness, attitudes, and actions related to COVID-19 among adults with chronic conditions at the onset of the U.S. outbreak: a cross-sectional survey. *Ann Intern Med* 2020 Jul 21;173(2):100-109 [FREE Full text] [doi: [10.7326/M20-1239](https://doi.org/10.7326/M20-1239)] [Medline: [32271861](https://pubmed.ncbi.nlm.nih.gov/32271861/)]
26. Murman D. The impact of age on cognition. *Semin Hear* 2015 Aug 9;36(3):111-121 [FREE Full text] [doi: [10.1055/s-0035-1555115](https://doi.org/10.1055/s-0035-1555115)] [Medline: [27516712](https://pubmed.ncbi.nlm.nih.gov/27516712/)]
27. Knowles B, Hanson V. The wisdom of older technology (non)users. *Commun ACM* 2018 Feb 21;61(3):72-77. [doi: [10.1145/3179995](https://doi.org/10.1145/3179995)]
28. Why some older people are rejecting digital technologies. *Science Daily*. 2018 Mar 12. URL: <https://www.sciencedaily.com/releases/2018/03/180312091715.htm>
29. Shawon A. Coronavirus: Bangladesh declares public holiday from March 26 to April 4. *Dhaka Tribune*. 2020 Mar 23. URL: <https://www.dhakatribune.com/bangladesh/2020/03/23/govt-offices-to-remain-closed-till-april-4> [accessed 2020-04-18]
30. Shammi M, Bodrud-Doza M, Islam ARMT, Rahman MM. Strategic assessment of COVID-19 pandemic in Bangladesh: comparative lockdown scenario analysis, public perception, and management for sustainability. *Environ Dev Sustain* 2020 Jul 18:1-44 [FREE Full text] [doi: [10.1007/s10668-020-00867-y](https://doi.org/10.1007/s10668-020-00867-y)] [Medline: [32837281](https://pubmed.ncbi.nlm.nih.gov/32837281/)]
31. Biswas RK, Huq S, Afiaz A. Relaxed lockdown in Bangladesh during COVID-19: should economy outweigh health? *Int J Health Policy Manag* 2020 Nov 01;9(11):488-490 [FREE Full text] [doi: [10.34172/ijhpm.2020.98](https://doi.org/10.34172/ijhpm.2020.98)] [Medline: [32610806](https://pubmed.ncbi.nlm.nih.gov/32610806/)]
32. General holidays likely to be further extended till May 16. *The Business Standard*. 2020 May 02. URL: <https://tbsnews.net/coronavirus-chronicle/covid-19-bangladesh/govt-extend-general-holidays-till-may-16-76291> [accessed 2021-10-18]
33. Huq S, Biswas RK. COVID-19 in Bangladesh: data deficiency to delayed decision. *J Glob Health* 2020 Jun;10(1):010342 [FREE Full text] [doi: [10.7189/jogh.10.010342](https://doi.org/10.7189/jogh.10.010342)] [Medline: [32373321](https://pubmed.ncbi.nlm.nih.gov/32373321/)]
34. Zhu J, Cai Y. Engaging the communities in Wuhan, China during the COVID-19 outbreak. *Glob Health Res Policy* 2020;5:35-36 [FREE Full text] [doi: [10.1186/s41256-020-00162-3](https://doi.org/10.1186/s41256-020-00162-3)] [Medline: [32685692](https://pubmed.ncbi.nlm.nih.gov/32685692/)]
35. Pollack T, Thwaites G, Rabaa M, Choisy M, van Doorn R, Tan LV, Tan, Exemplars in Global Health. Emerging COVID-19 success story: Vietnam's commitment to containment. *Our World in Data*. URL: <https://ourworldindata.org/covid-exemplar-vietnam> [accessed 2021-10-18]
36. Urban population (% of total population) - Bangladesh. World Bank. 2019. URL: <https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=BD> [accessed 2020-08-03]
37. Bangladesh Demographic and Health Survey 2017-2018: Key Indicators. Ministry of Health and Family Welfare, Dhaka, Bangladesh. 2019. URL: <https://dhsprogram.com/pubs/pdf/PR104/PR104.pdf> [accessed 2021-10-18]
38. COVID-19: Bangladesh MultiSectoral Anticipatory Impact and Needs Analysis. Needs Assessment Working Group Bangladesh. 2020 Apr 27. URL: https://reliefweb.int/sites/reliefweb.int/files/resources/covid_nawg_anticipatory_impacts_and_needs_analysis.pdf [accessed 2021-10-18]
39. IOM assists vulnerable returning migrants impacted by the COVID-19 pandemic. International Organization for Migration, Bangladesh. 2020 May 20. URL: <https://bangladesh.iom.int/news/iom-assists-vulnerable-returning-migrants-impacted-covid-19-pandemic> [accessed 2020-05-20]
40. Das S, Mia M, Hanifi S, Hoque S, Bhuiya A. Health literacy in a community with low levels of education: findings from Chakaria, a rural area of Bangladesh. *BMC Public Health* 2017 Feb 16;17(1):203 [FREE Full text] [doi: [10.1186/s12889-017-4097-y](https://doi.org/10.1186/s12889-017-4097-y)] [Medline: [28209185](https://pubmed.ncbi.nlm.nih.gov/28209185/)]
41. Zahnd W, Scaife SL, Francis ML. Health literacy skills in rural and urban populations. *Am J Health Behav* 2009;33(5):550-557. [doi: [10.5993/ajhb.33.5.8](https://doi.org/10.5993/ajhb.33.5.8)] [Medline: [19296745](https://pubmed.ncbi.nlm.nih.gov/19296745/)]

42. Azlan AA, Hamzah MR, Sern TJ, Ayub SH, Mohamad E. Public knowledge, attitudes and practices towards COVID-19: a cross-sectional study in Malaysia. PLoS One 2020;15(5):e0233668 [FREE Full text] [doi: [10.1371/journal.pone.0233668](https://doi.org/10.1371/journal.pone.0233668)] [Medline: [32437434](https://pubmed.ncbi.nlm.nih.gov/32437434/)]
43. Reuben RC, Danladi MMA, Saleh DA, Ejembi PE. Knowledge, attitudes and practices towards COVID-19: an epidemiological survey in North-Central Nigeria. J Community Health 2021 Jun;46(3):457-470 [FREE Full text] [doi: [10.1007/s10900-020-00881-1](https://doi.org/10.1007/s10900-020-00881-1)] [Medline: [32638198](https://pubmed.ncbi.nlm.nih.gov/32638198/)]
44. Ryan P. Integrated Theory of Health Behavior Change: background and intervention development. Clin Nurse Spec 2009;23(3):161-70; quiz 171 [FREE Full text] [doi: [10.1097/NUR.0b013e3181a42373](https://doi.org/10.1097/NUR.0b013e3181a42373)] [Medline: [19395894](https://pubmed.ncbi.nlm.nih.gov/19395894/)]
45. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: [10.1146/annurev.publhealth.012809.103604](https://doi.org/10.1146/annurev.publhealth.012809.103604)] [Medline: [20070207](https://pubmed.ncbi.nlm.nih.gov/20070207/)]
46. Risk communication and community engagement (RCCE) action plan guidance: COVID-19 preparedness and response. World Health Organization. 2020 Mar 16. URL: <https://www.who.int/publications/i/item/risk-communication-and-community-engagement-> [accessed 2020-09-12]

Abbreviations

KAP: knowledge, attitudes, and practices

OR: odds ratio

WHO: World Health Organization

Edited by G Eysenbach; submitted 02.03.21; peer-reviewed by M Al-Hanawi, R Shahriarirad; comments to author 14.03.21; revised version received 04.04.21; accepted 22.04.21; published 05.11.21.

Please cite as:

Rabbani MG, Akter O, Hasan MZ, Samad N, Mahmood SS, Joarder T

COVID-19 Knowledge, Attitudes, and Practices Among People in Bangladesh: Telephone-Based Cross-sectional Survey

JMIR Form Res 2021;5(11):e28344

URL: <https://formative.jmir.org/2021/11/e28344>

doi: [10.2196/28344](https://doi.org/10.2196/28344)

PMID: [34519660](https://pubmed.ncbi.nlm.nih.gov/34519660/)

©Md Golam Rabbani, Orin Akter, Md Zahid Hasan, Nandeeta Samad, Shehrin Shaila Mahmood, Taufique Joarder. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.11.2021. 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

Exploring Online Health Reviews to Monitor COVID-19 Public Health Responses in Alabama State Department of Corrections: Case Example

Pamela Valera¹, MSW, PhD; David Carmona¹, BS; Sarah Malarkey¹, BS; Noah Sinangil¹, BS; Madelyn Owens¹, BA; Asia Lefebvre¹, BA

Department of Urban-Global Public Health, Rutgers School of Public Health, Newark, NJ, United States

Corresponding Author:

Pamela Valera, MSW, PhD

Department of Urban-Global Public Health

Rutgers School of Public Health

1 Riverfront Plaza, 10th Floor

Newark, NJ, 07102

United States

Phone: 1 9739726555

Email: pv181@sph.rutgers.edu

Abstract

Background: COVID-19, caused by SARS-CoV-2, has devastated incarcerated people throughout the United States.

Objective: The purpose of this study was to test the feasibility and acceptability of a COVID-19 Health Review for Correctional Facilities.

Methods: The COVID-19 Health Review survey for the Department of Corrections was developed in Qualtrics to assess the following: (1) COVID-19 testing, (2) providing personal protective equipment, (3) vaccination procedures, (4) quarantine procedures, (5) COVID-19 mortality rates for inmates, (6) COVID-19 mortality rates for correctional officers and prison staff, (7) COVID-19 infection rates for inmates, (8) COVID-19 infection rates for correctional officers and prison staff, and (9) uptake of COVID-19 vaccines. The estimated time to review the Alabama State Department of Corrections COVID-19 responses on their website and complete the survey items was 45 minutes to 1 hour.

Results: Of the 21 participants who completed the COVID-19 Health Review for Correctional Facilities survey, 48% (n=10) identified as female, 43% (n=9) identified as male, and 10% (n=2) identified as transgender. For race, 29% (n=6) self-identified as Black or African American, 24% (n=5) Asian, 24% (n=5) White, 5% (n=1) Pacific Islander or Native Hawaiian, and 19% (n=4) Other. In addition, 5 respondents self-identified as returning citizens. For COVID-19 review questions, the majority concluded that information on personal protective equipment was “poor” and “very poor,” information on COVID-19 testing was “fair” and above, information on COVID-19 death/infection rates between inmates and staff was “good” and “very good,” and information on vaccinations was “good” and “very good.” There was a significant difference observed ($P=.03$) between nonreturning citizens and returning citizens regarding the health grade review with respect to available information on COVID-19 infection rates.

Conclusions: COVID-19 health reviews may provide an opportunity for the public to review the COVID-19 responses in correctional settings.

(*JMIR Form Res* 2021;5(11):e32591) doi:[10.2196/32591](https://doi.org/10.2196/32591)

KEYWORDS

Alabama; correctional facilities; COVID-19; online health reviews; review; monitoring; public health; policy; response; prison; United States; case study; formative; feasibility; acceptability; survey

Introduction

Background

COVID-19, caused by SARS-CoV-2, has devastated people nationwide in the United States, particularly people who are incarcerated. There are 2.3 million incarcerated people in the United States that are imprisoned in various correctional institutions across the criminal justice system [1]. As of June 2021, nearly 400,000 infectious cases and close to 3000 deaths in prison were attributed to COVID-19 [2,3]. Furthermore, COVID-19 infection rates in prisons were found to be 5.5 times higher than that of the general US population [4], and mortality rates due to COVID-19 in correctional facilities were found to be 2.5 times higher than that of the general US population [4,5]. COVID-19 outbreaks across correctional settings, including Ohio's Cook County Jail accounting for 15% of Chicago's total COVID-19 cases [6] and Alabama's COVID-19 outbreak in Ashville Jail with 37 inmates testing positive [7], exposed the deep structural and administrative problems with limited public health measures in place. The COVID-19 preventative measures of social distancing and quarantining are nearly impossible to implement as incarcerated people share common areas such as cell units, bathroom stalls, showers, and cafeterias, among others. These living conditions are often overcrowded and poorly ventilated [8-11].

Concurrently, correctional officers and other prison staff workers are susceptible to COVID-19 as they are in and out of facilities, which increases the likelihood of spreading and transmitting the virus [9,10,12,13]. The COVID-19 pandemic has further reduced the already declining number of correctional staff, severely impacting public health responses [13-15].

The incarcerated population is also aging and typically older; at least half are 50 years or older, report having a chronic condition, and are 2.5 times more likely to report a chronic illness than the general population [16]. COVID-19 typically has a greater impact on older people, particularly those with preexisting conditions [17]. The lack of inadequate and unequal access to health care services may further exacerbate the impact of COVID-19 on correctional settings and its incarcerated people [8,13].

Understanding the State Department of Corrections' Response to COVID-19

The State Department of Corrections' response to COVID-19 was delayed during the early days of the pandemic due to the novelty of the disease [18]. When the Centers for Disease Control and Prevention (CDC) called for face masks and gloves to be worn when social distancing was impossible, the US Bureau of Prisons (BOP) confirmed that personal protective equipment (PPE) had been provided to all inmates and prison staff [19]. Yet, reports of unavailable PPE in various prisons and the national shortage of PPE resulted in the failed initial response to COVID-19 in prisons [19].

BOP Rules and Regulations Concerning Inmate Transfers

The BOP stated that inmates could be transferred if they met three conditions: (1) passing a COVID-19 exit screening or not showing any symptoms, (2) being held for 14 days in BOP custody, and (3) notifying the BOP Emergency Operations Center before the transfer of an inmate [20]. Although prison visitations were prohibited, many incarcerated people continued to be transferred without proper COVID-19 testing [18]. Indeed, testing and contact tracing were relatively slow, causing infection cases to go undetected [18,21]. One of the more sought-after responses to combat COVID-19 was to reduce jail and prison populations by releasing incarcerated people, which would have slowed the spread and death of the prison population [22-24]. However, during the latter half of 2020, many prison systems such as those in Wyoming, Mississippi, Montana, and Hawaii were still at 90% or higher capacity [25].

Online Health Reviews

Understanding how the US State Department of Corrections responded to the COVID-19 pandemic in prisons may help to build best practices regarding the detection and prevention of COVID-19. Hence, it is vital to evaluate and assess actions taken as part of a public health response to reduce the spread of COVID-19 in prisons. One public health approach that has been shown to connect people with vital information about local organizations is online health reviews.

Online health reviews are designed to help consumers find vetted local providers such as dentists, physicians, other specialty care providers, and many more. Online health reviews, or "health grades," are generally used to review the quality of a physician's or provider's care or a patient's visit. Over the past few years, there has been significant growth in the use and development of health reviews and ratings in the United States [26]. Online reviews such as Yelp or HealthGrades may provide the public an opportunity to monitor these venues and improve the state or local inspection efforts. Consumers who review their experience and leave reviews help others make informed decisions about the quality of care they may receive from the specific hospital or provider. In addition, studies have shown that online health reviews can help surgeons, physicians, and other medical providers improve their overall care and treatment of patients [26-28]. Most importantly, the increased usage of online reviews can also impact patient choice. Reviews that express a negative comment can influence another patient's decision to seek medical care at that location [27].

Online Reviews in Providing Patient Choices

Online health reviews may provide people with confidence that they are making informed choices by reading through the reviews about the business and helping to recommend their provider to friends and family [29]. When we look deeper into the content of reviews, patients often talk about relationships with their medical provider, suggesting that such interactions may be essential to a patient when leaving a review [30]. Health reviews undoubtedly play a vital role in sharing patient experience and even influencing patient choice. Although online health reviews may play a critical role in influencing a patient's

choice, reviews about correctional facilities' public health response to COVID-19 are relatively unknown. Online health reviews could potentially improve COVID-19 responses in correctional facilities [31].

Objective of the Present Study

Given the urgent need to prevent and reduce the spread of COVID-19 in correctional facilities, this project aims to further understand the COVID-19 response of one State Department of Corrections by developing a health review survey using the Alabama State Department of Corrections (ADOC) as a case study. The ADOC was selected because it has one of the highest COVID-19 mortality rates among incarcerated people in the United States [2,3] and has faced substantial challenges in providing adequate public health responses to the COVID-19 pandemic [32,33].

The ADOC has had 2085 positive cases and 68 fatalities from COVID-19 among inmates [34]. COVID-19 vaccinations have been administered to 48.06% (11,895/24,751) of incarcerated people housed in their facilities [34]. The ADOC's website provides detailed information on preventative and protective measures, testing updates, and operation plans. However, the ADOC has failed to provide proper COVID-19 testing and PPE, reduce inmate populations, release medically vulnerable individuals and those near the end of their sentences, and publicly make data available on COVID-19 [31]. Moreover,

the ADOC has been fraught with controversy in recent months for denying COVID-19 vaccines to its incarcerated people. Litigation is underway by the US Justice Department against the ADOC for unsanitary and disturbing trends in reports of violence that exacerbate their inadequate response to COVID-19 [32,33].

Methods

Survey

The COVID-19 Health Review survey for the Department of Corrections was developed in Qualtrics to assess the following: (1) COVID-19 testing, (2) providing PPE, (3) vaccination procedures, (4) quarantine procedures, (5) COVID-19 mortality rates for inmates, (6) COVID-19 mortality rates for correctional officers and prison staff, (7) COVID-19 infection rates for inmates, (8) COVID-19 infection rates for correctional officers and prison staff, and (9) uptake of COVID-19 vaccines. To better understand the information on the Department of Corrections websites, the survey included eight items that focused on the CDC's public health measures. [Textbox 1](#) provides the list of questions. Participants were then asked to respond to each item on a 6-point scale ("very poor," "poor," "fair," "good," "very good," or "unknown"). The survey was anonymous and no personal information was collected. The estimated completion time of the survey ranged from 45 minutes to 1 hour.

Textbox 1. Questions related to public safety items for COVID-19 measures on the website.

ALABAMA: Please review the Department of Corrections (DOC) website and rank the information for each component. To access the DOC website, right-click "Alabama" and select "open link in a new tab."

1. Does the website showcase information about providing protective equipment to incarcerated persons?
2. Is there policy in place for the prison to engage with public health recommendations like social distancing?
3. Are there data that inform the public about COVID-19 death rates among incarcerated persons?
4. Are there data that inform the public about COVID-19 death rates for correctional officers and prison staff?
5. Are there data that describe COVID-19 infection rates for incarcerated persons?
6. Are there data that describe COVID-19 infection rates for correctional officers and prison staff?
7. Has anyone been vaccinated within the prison setting?

Note: Each of these items was set on a 6-point scale: "very poor," "poor," "fair," "good," "very good," or "unknown"

Recruitment of Survey Participants

Participants were recruited via social media using email from a Listserv, Twitter, and Instagram from April 2021 to July 2021. A total of 35 participants completed the survey.

Statistical Analysis

Demographic characteristics and COVID-19 health review results are reported with frequencies and percentages. The results were then stratified by those who self-reported that they were returning citizens or people with a criminal justice history and those who were not returning citizens to identify any significant

differences. Fisher exact tests [35] were performed for this analysis, as the data were not normally distributed.

Results

Thirty-five participants completed the survey; after removing 14 observations due to incomplete data, 21 observations remained for final analyses. Of the 21 participants, five identified as returning citizens and 16 did not report criminal justice involvement. The basic characteristics of the participants and comparisons between returning and nonreturning citizens are summarized in [Table 1](#).

Table 1. Demographic characteristics of the participants.

Variable	Total sample (N=21), n (%) ^a	Returning citizens (n=5), n (%)	Nonreturning citizens (n=16), n (%) ^a	<i>P</i> value ^b
Age group (years)				.07
18-24	8 (38)	1 (20)	7 (44)	
25-34	4 (19)	0 (0)	4 (25)	
35-44	4 (19)	2 (40)	2 (13)	
45-54	2 (10)	0 (0)	2 (13)	
55-64	2 (10)	2 (40)	0 (0)	
≥65	1 (5)	0 (0)	1 (6)	
Gender				.79
Female	10 (48)	2 (40)	8 (50)	
Male	9 (43)	3 (60)	6 (38)	
Transgender	2 (10)	0 (0)	2 (13)	
Race (check all that apply)				.08
White	5 (24)	0 (0)	5 (31)	
Black	4 (19)	1 (20)	3 (19)	
African American	2 (10)	1 (20)	1 (6)	
Asian	5 (24)	0 (0)	5 (31)	
Pacific Islander/Native Hawaiian	1 (5)	1 (20)	0 (0)	
Other	4 (19)	2 (40)	2 (13)	
Hispanic or Latin origin				.55
Yes	5 (24)	2 (40)	3 (19)	
No	16 (76)	3 (60)	13 (81)	
Highest level of education achieved				.03
High school	8 (38)	4 (80)	4 (25)	
Bachelor's degree	9 (43)	0 (0)	9 (56)	
Master's degree	2 (10)	0 (0)	2 (13)	
PhD or higher	1 (5)	0 (0)	1 (6)	
Trade school	1 (5)	1 (20)	0 (0)	
Current employment status				.12
Employed full time	9 (43)	4 (80)	5 (31)	
Employed part time	8 (38)	0 (0)	8 (50)	
Seeking opportunities	4 (19)	1 (20)	3 (19)	

^aDue to rounding, percentages may not add up to 100.

^bFisher exact test.

Of the 21 participants, about 40% (n=8) were 18-24 years old. The sample was racially diverse, with about 30% (n=6) being Black or African American, a quarter White, followed by Asian Americans, Pacific Islander or Native Hawaiian, and a sizable "other" race category. Nearly half of the participants reported their gender identity as female, followed by male, and two individuals identified as transgender. The majority (76%, n=16) of the sample did not identify as being of Hispanic or Latino origin. Nearly 80% (n=17) of the participants were either

employed part time or full time. The highest level of education was the only significant association between returning citizens and nonreturning citizens ($P=.03$). Eighty percent (n=4) of the returning citizens' highest degree of education achieved was high school compared to the majority (75%, n=12) of the nonreturning citizens earning a bachelor's degree or higher. [Table 2](#) compares the results of the COVID-19 Health Review items for the ADOC between returning and nonreturning citizens.

Table 2. Comparative COVID-19 grade results among returning citizens and nonreturning citizens.

Health Review item	Total sample (N=21), n (%) ^a	Returning citizens (n=5), n (%)	Nonreturning citizens (n=16), n (%) ^a	<i>P</i> value ^b
Does the website provide information relating to COVID-19 testing?				.94
Very poor	2 (10)	1 (20)	1 (6)	
Poor	3 (14)	1 (20)	2 (13)	
Fair	5 (24)	1 (20)	4 (25)	
Good	6 (29)	1 (20)	5 (31)	
Very good	5 (24)	1 (20)	4 (25)	
Unknown	0 (0)	0 (0)	0 (0)	
Does the website showcase information about providing protective equipment to incarcerated persons?				.16
Very poor	4 (19)	2 (40)	2 (13)	
Poor	7 (33)	1 (20)	6 (38)	
Fair	1 (5)	1 (20)	0 (0)	
Good	3 (14)	0 (0)	3 (19)	
Very good	2 (10)	1 (20)	1 (6)	
Unknown	4 (19)	0 (0)	4 (25)	
Is there policy in place to facilitate vaccination for all incarcerated persons and prison staff?				.69
Very poor	3 (14)	1 (20)	2 (13)	
Poor	3 (14)	1 (20)	2 (13)	
Fair	6 (29)	2 (40)	4 (25)	
Good	4 (19)	0 (0)	4 (25)	
Very good	2 (10)	1 (20)	1 (6)	
Unknown	3 (14)	0 (0)	3 (19)	
Is there policy in place for the prison to engage with public health recommendations like social distancing?				.29
Very poor	3 (14)	1 (20)	2 (13)	
Poor	7 (33)	1 (20)	6 (38)	
Fair	3 (14)	2 (40)	1 (6)	
Good	2 (10)	1 (20)	1 (6)	
Very good	2 (10)	0 (0)	2 (13)	
Unknown	4 (19)	0 (0)	4 (25)	
Are there data that inform the public about COVID-19 death rates for incarcerated persons?				.51
Very poor	1 (5)	0 (0)	1 (6)	
Poor	3 (14)	2 (40)	1 (6)	
Fair	2 (10)	0 (0)	2 (13)	
Good	6 (29)	1 (20)	5 (31)	
Very good	9 (43)	2 (40)	7 (44)	
Unknown	0 (0)	0 (0)	0 (0)	
Are there data that inform the public about COVID-19 death rates for correctional officers and prison staff?				.09
Very poor	1 (5)	0 (0)	1 (6)	
Poor	3 (14)	2 (40)	1 (6)	

Health Review item	Total sample (N=21), n (%) ^a	Returning citizens (n=5), n (%)	Nonreturning citizens (n=16), n (%) ^a	<i>P</i> value ^b
Fair	3 (14)	2 (40)	1 (6)	
Good	8 (38)	1 (20)	7 (44)	
Very good	5 (24)	0 (0)	5 (31)	
Unknown	1 (5)	0 (0)	1 (6)	
Are there data that describe COVID-19 infection rates for incarcerated persons?				.03
Very poor	1 (5)	1 (20)	0 (0)	
Poor	4 (19)	2 (40)	2 (13)	
Fair	1 (5)	1 (20)	0 (0)	
Good	8 (38)	0 (0)	8 (50)	
Very good	6 (29)	1 (20)	5 (31)	
Unknown	1 (5)	0 (0)	1 (6)	
Are there data that describe COVID-19 infection rates for correctional officers and prison staff?				.07
Very poor	1 (5)	1 (20)	0 (0)	
Poor	4 (19)	2 (40)	2 (13)	
Fair	1 (5)	1 (20)	0 (0)	
Good	9 (43)	1 (20)	8 (50)	
Very good	4 (19)	0 (0)	4 (25)	
Unknown	2 (10)	0 (0)	2 (13)	
Uptake of COVID-19 vaccines: has anyone been vaccinated within the prison setting?				.30
Very poor	0 (0)	0 (0)	0 (0)	
Poor	3 (14)	2 (40)	1 (6)	
Fair	2 (10)	0 (0)	2 (13)	
Good	7 (33)	1 (20)	6 (38)	
Very good	7 (33)	1 (20)	6 (38)	
Unknown	2 (10)	1 (20)	1 (6)	

^aDue to rounding, percentages may not add up to 100.

^bFisher exact test.

Sixteen respondents reported that the information provided on the website related to COVID-19 testing was “fair” and above. Despite having “fair” information about COVID-19 testing, half of the respondents (n=11) reported the details on providing PPE to incarcerated persons as “poor” and “very poor.” Furthermore, approximately half (n=12) of the respondents reported the information about policies to facilitate vaccinations for all incarcerated persons and prison staff as “fair” and above. Approximately 10 of the respondents reported guidelines for the prison to engage with public health recommendations as “very poor” and “poor.” Approximately 15 participants of the sample reported information on COVID-19 death rates among incarcerated persons as “good” and “very good.” A total of 13 participants provided the same rating with respect to data about correctional officers and prison staff.

With regard to COVID-19 infection rates, most respondents reported that the ADOC’s COVID-19 website provided “good” and “very good” data for both incarcerated persons and correctional officers. Two-thirds (n=14) of the sample reported that the information provided about whether people were vaccinated within the prison setting was “good” and “very good.” The only significant difference between returning and nonreturning citizens was whether the data included on the ADOC website describe incarcerated people’s COVID-19 infection rates ($P=0.03$): 80% (n=4) of the returning citizens reported the data included as “fair” and below compared to 75% (n=12) of the nonreturning citizens reporting the data included as “good” and “very good.”

Participants were also asked to provide feedback about the ADOC COVID-19 website. Example quotes of some of this feedback are provided in [Textbox 2](#).

Textbox 2. Representative feedback about the Alabama Department of Corrections' COVID-19 website.

"They had the basic info. The more in-depth information was difficult to find."

"No information on COVID-19 procedures."

"I quickly found all other information. This took me about 3 minutes. The COVID information was very clear, but in English only."

"The website did not provide information on PPE [personal protective equipment] equipment and how vaccines will be given to inmates. COVID-19 preparedness is shown in a single document that describes everything but specific quarantine procedures. The website had a dashboard, so it was fairly quick to find basic information on COVID-19 such as rates."

"The information is there, but little. And, it seems that they are going to start the vaccine. I looked a lot because I was looking for information about PPE."

Discussion

Principal Findings

This online COVID-19 public health review in a correctional setting provides an opportunity to examine the current operational practices at the population level in response to the COVID-19 pandemic. This is the first study to explore the role of digital health response to outbreaks, vaccines, and pandemics that focuses on correctional facilities, which may help to reduce the deadly impact of COVID-19 on the lives of marginalized communities. Our findings build on emerging literature that describes online health review topics that could potentially be applied in public health responses to COVID-19 [29,36,37].

We believe that these results may serve as an essential reference point for policymakers and advocates to understand the impact and relevance of COVID-19 online health reviews, and to prioritize resources and efforts to address the challenges presented by COVID-19 in correctional settings. For example, most participants felt that the reported PPE information provided on the ADOC website was "poor" and "very poor." This implies that information provided to the public about PPE may be limited. At the same time, there was a significant difference ($P=.03$) observed between nonreturning citizens and returning citizens on the quality of information about COVID-19 infection cases displayed on the ADOC website. The majority of returning citizens found this to be "fair." In contrast, the participants without a criminal justice history found the information to be "good" and "very good." Although participants noted that COVID-19 public health measures were available on the ADOC website, this information was only available in English. An official website of the US government should be available in multiple languages.

Furthermore, using health grades or online reviews to understand the mitigation and adaptation responses to COVID-19 is critical to prevent the infection and mortality rates in correctional settings. This would allow for additional oversight and place pressure on the State Department of Corrections to do the right thing, especially during a pandemic where correctional officers and incarcerated persons are directly impacted. As previously mentioned, there has been no use of health grades or online reviews in prison systems to date.

Our study presents an opportunity to enter a new landscape of grading or reviewing correctional facility responses toward COVID-19. The existing scientific literature on health grades and online reviews has stressed how reviews can influence people's perceptions and choices regarding seeking care and even recommending care [27,29] while also having the potential to improve the quality of care and treatment [26-28]. Our work builds upon the current literature by extending the potential of influencing people's choices and perceptions toward how correctional facilities are handling COVID-19 and improving these same correctional facilities to better respond to the impact of COVID-19.

Implications for Future Research

Given how critical online reviews can be, there is much potential in using online reviews to cover more aspects of correctional facilities. For example, although we specifically focused on the ADOC in this study, there may be potential to create a nationwide health grading system covering every Department of Corrections. For example, the web-based company HealthGrades has a database on all doctors in the United States and includes patient reviews for every doctor [38]. In a similar vein, we may create an online database that keeps track of all COVID-19 responses across all Departments of Corrections in the United States. This information is vital for frontline workers at correctional facilities to improve and provide a safer and secure service for incarcerated people.

Limitations

Our survey results have several limitations that must be acknowledged. First, not all recruited participants were able to complete the COVID-19 correctional health survey, resulting in incomplete and missing data. The length of the survey is also a limitation, as the 1-hour completion time may have impacted our participants' ability to examine the ADOC website properly.

Conclusions

The COVID-19 pandemic is still ongoing and new technologies are needed to evaluate the impact of COVID-19 in correctional settings. There is added value to examine how online health reviews can be used to understand the COVID-19 pandemic response in correctional facilities. Online health reviews offer tools to inform the public about how the Department of Corrections supports a public health pandemic response that saves lives.

Acknowledgments

We thank the Community Health Justice Lab members for helping with the recruitment of potential participants to complete the survey.

Conflicts of Interest

None declared.

References

1. Sawyer W, Wagner P. Mass incarceration: the whole pie 2020. Prison Policy Initiative. URL: <https://www.prisonpolicy.org/reports/pie2020.html> [accessed 2021-07-30]
2. COVID-19. The Marshall Project. 2021 Oct 27. URL: <https://www.themarshallproject.org/records/8793-covid-19> [accessed 2021-07-30]
3. The COVID Prison Project tracks data and policy across the country to monitor COVID-19 in prisons. COVID Prison Project. URL: <https://covidprisonproject.com/> [accessed 2021-07-30]
4. Saloner B, Parish K, Ward JA, DiLaura G, Dolovich S. COVID-19 cases and deaths in federal and state prisons. JAMA 2020 Aug 11;324(6):602-603 [FREE Full text] [doi: [10.1001/jama.2020.12528](https://doi.org/10.1001/jama.2020.12528)] [Medline: [32639537](https://pubmed.ncbi.nlm.nih.gov/32639537/)]
5. Toblin RL, Hagan LM. COVID-19 case and mortality rates in the Federal Bureau of Prisons. Am J Prev Med 2021 Jul;61(1):120-123 [FREE Full text] [doi: [10.1016/j.amepre.2021.01.019](https://doi.org/10.1016/j.amepre.2021.01.019)] [Medline: [33781619](https://pubmed.ncbi.nlm.nih.gov/33781619/)]
6. Reinhart E, Chen DL. Incarceration and its disseminations: COVID-19 pandemic lessons from Chicago's Cook County Jail. Health Aff (Millwood) 2020 Aug 01;39(8):1412-1418. [doi: [10.1377/hlthaff.2020.00652](https://doi.org/10.1377/hlthaff.2020.00652)] [Medline: [32496864](https://pubmed.ncbi.nlm.nih.gov/32496864/)]
7. The St. Clair County Jail in Ashville is Currently Experiencing a COVID 19 Outbreak. Clair County Alabama Facebook. URL: <https://www.facebook.com/StClairCountyAlabamaSheriff/posts/189489866524869> [accessed 2021-07-29]
8. Hawks L, Woolhandler S, McCormick D. COVID-19 in prisons and jails in the United States. JAMA Intern Med 2020 Aug 01;180(8):1041-1042. [doi: [10.1001/jamainternmed.2020.1856](https://doi.org/10.1001/jamainternmed.2020.1856)] [Medline: [32343355](https://pubmed.ncbi.nlm.nih.gov/32343355/)]
9. Denied vaccines, people incarcerated in Alabama prisons are dying of Covid. Equal Justice Initiative. URL: <https://eji.org/news/denied-vaccines-people-in-alabama-prisons-are-dying-of-covid/> [accessed 2021-07-29]
10. Kinner SA, Young JT, Snow K, Southalan L, Lopez-Acuña D, Ferreira-Borges C, et al. Prisons and custodial settings are part of a comprehensive response to COVID-19. Lancet Public Health 2020 Apr;5(4):e188-e189 [FREE Full text] [doi: [10.1016/S2468-2667\(20\)30058-X](https://doi.org/10.1016/S2468-2667(20)30058-X)] [Medline: [32197116](https://pubmed.ncbi.nlm.nih.gov/32197116/)]
11. Franco-Paredes C, Jankowsky K, Schultz J, Bernfeld J, Cullen K, Quan NG, et al. COVID-19 in jails and prisons: a neglected infection in a marginalized population. PLoS Negl Trop Dis 2020 Jun 22;14(6):e0008409-e0008409 [FREE Full text] [doi: [10.1371/journal.pntd.0008409](https://doi.org/10.1371/journal.pntd.0008409)] [Medline: [32569274](https://pubmed.ncbi.nlm.nih.gov/32569274/)]
12. Rubin R. The challenge of preventing COVID-19 spread in correctional facilities. JAMA 2020 May 12;323(18):1760-1761. [doi: [10.1001/jama.2020.5427](https://doi.org/10.1001/jama.2020.5427)] [Medline: [32259189](https://pubmed.ncbi.nlm.nih.gov/32259189/)]
13. Montoya-Barthelemy AG, Lee CD, Cundiff DR, Smith EB. COVID-19 and the correctional environment: the American prison as a focal point for public health. Am J Prev Med 2020 Jun;58(6):888-891 [FREE Full text] [doi: [10.1016/j.amepre.2020.04.001](https://doi.org/10.1016/j.amepre.2020.04.001)] [Medline: [32387174](https://pubmed.ncbi.nlm.nih.gov/32387174/)]
14. Barnert E, Ahalt C, Williams B. Prisons: amplifiers of the COVID-19 pandemic hiding in plain sight. Am J Public Health 2020 Jul;110(7):964-966. [doi: [10.2105/AJPH.2020.305713](https://doi.org/10.2105/AJPH.2020.305713)] [Medline: [32407126](https://pubmed.ncbi.nlm.nih.gov/32407126/)]
15. Derr B, Griesba R, Issawi D. States are shutting down prisons as guards are crippled By Covid-19. The New York Times. URL: <https://www.nytimes.com/2021/01/01/us/coronavirus-prisons-jails-closing.html> [accessed 2021-07-30]
16. Maruschak L, Berzofsky M, Unangst J. Medical problems of state and federal prisoners and jail inmates, 2011-12. Bureau of Justice Statistics. 2016 Oct 4. URL: <https://bjs.ojp.gov/content/pub/pdf/mpsfj1112.pdf> [accessed 2021-10-27]
17. COVID-19 and your health: People with certain medical conditions. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html> [accessed 2021-07-29]
18. Burkhalter E, Colón I, Derr B. Incarcerated and infected: how the virus tore through the U.S. prison system. The New York Times. URL: <https://www.nytimes.com/interactive/2021/04/10/us/covid-prison-outbreak.html> [accessed 2021-07-29]
19. Hummer D. United States Bureau of Prisons' response to the COVID-19 pandemic. Victims Offend 2020 Oct 22;15(7-8):1262-1276. [doi: [10.1080/15564886.2020.1829765](https://doi.org/10.1080/15564886.2020.1829765)]
20. Marcum CD. American Corrections System response to COVID-19: an examination of the procedures and policies used in spring 2020. Am J Crim Justice 2020 Jun 07;45(4):759 [FREE Full text] [doi: [10.1007/s12103-020-09535-3](https://doi.org/10.1007/s12103-020-09535-3)] [Medline: [32837155](https://pubmed.ncbi.nlm.nih.gov/32837155/)]
21. Nelson B, Kaminsky DB. A COVID-19 crisis in US jails and prisons. Cancer Cytopathol 2020 Aug 03;128(8):513-514. [doi: [10.1002/ency.22335](https://doi.org/10.1002/ency.22335)] [Medline: [32744804](https://pubmed.ncbi.nlm.nih.gov/32744804/)]
22. Macmadu A, Berk J, Kaplowitz E, Mercedes M, Rich JD, Brinkley-Rubinstein L. COVID-19 and mass incarceration: a call for urgent action. Lancet Public Health 2020 Nov;5(11):e571-e572 [FREE Full text] [doi: [10.1016/S2468-2667\(20\)30231-0](https://doi.org/10.1016/S2468-2667(20)30231-0)] [Medline: [33045186](https://pubmed.ncbi.nlm.nih.gov/33045186/)]

23. Herring T, Sharma M. States of emergency: the failure of prison system responses to COVID-19. Prison Policy Institute. URL: https://www.prisonpolicy.org/reports/states_of_emergency.html [accessed 2021-09-01]
24. Burki T. Prisons are "in no way equipped" to deal with COVID-19. *Lancet* 2020 May 02;395(10234):1411-1412 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30984-3](https://doi.org/10.1016/S0140-6736(20)30984-3)] [Medline: [32359457](https://pubmed.ncbi.nlm.nih.gov/32359457/)]
25. Kajstura A, Landon J. Since you asked: is social distancing possible behind bars? Prison Policy Institute. 2020 Apr 03. URL: <https://www.prisonpolicy.org/blog/2020/04/03/density/> [accessed 2007-07-29]
26. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res* 2012 Feb 24;14(1):e38 [FREE Full text] [doi: [10.2196/jmir.2003](https://doi.org/10.2196/jmir.2003)] [Medline: [22366336](https://pubmed.ncbi.nlm.nih.gov/22366336/)]
27. Furnas H, Korman J, Canales F, Pence L. Patient reviews: Yelp, Google, Healthgrades, Vitals, and RealSelf. *Plast Reconstr Surg* 2020 Dec;146(6):1419-1431. [doi: [10.1097/PRS.00000000000007379](https://doi.org/10.1097/PRS.00000000000007379)] [Medline: [33234982](https://pubmed.ncbi.nlm.nih.gov/33234982/)]
28. Lin Y, Hong YA, Henson BS, Stevenson RD, Hong S, Lyu T, et al. Assessing patient experience and healthcare quality of dental care using patient online reviews in the United States: mixed methods study. *J Med Internet Res* 2020 Jul 07;22(7):e18652 [FREE Full text] [doi: [10.2196/18652](https://doi.org/10.2196/18652)] [Medline: [32673240](https://pubmed.ncbi.nlm.nih.gov/32673240/)]
29. Hong YA, Liang C, Radcliff TA, Wigfall LT, Street RL. What do patients say about doctors online? A systematic review of studies on patient online reviews. *J Med Internet Res* 2019 Apr 8;21(4):e12521. [doi: [10.2196/12521](https://doi.org/10.2196/12521)]
30. He L, He C, Wang Y, Hu Z, Zheng K, Chen Y. What do patients care about? Mining fine-grained patient concerns from online physician reviews through computer-assisted multi-level qualitative analysis. In: AMIA Annu Symp Proc. 2021 Jan 25 Presented at: AMIA 2020 Virtual Annual Symposium; November 14-18, 2020; Virtual (online) p. 544-553.
31. Widra E, Hayre D. Prison Policy Institute. URL: https://www.prisonpolicy.org/reports/failing_grades.html [accessed 2021-07-29]
32. Diaz J. Justice Department sues Alabama over prison conditions. National Public Radio (NPR). 2020 Dec 10. URL: <https://www.npr.org/2020/12/10/944879818/justice-dept-sues-alabama-over-prison-conditions> [accessed 2021-07-29]
33. Covid-19's impact on people in prison. Equal Justice Initiative. URL: <https://eji.org/news/covid-19s-impact-on-people-in-prison/> [accessed 2021-07-29]
34. Communication with population. Alabama Department of Corrections. URL: <http://www.doc.alabama.gov/COVID19Communication> [accessed 2021-10-01]
35. Fisher RA. *Statistical Methods for Research Workers* (5th Ed). Edinburgh: Oliver and Boyd; Jan 01, 1935:76.
36. Kemp E, Porter III M, Albert C, Min KS. Information transparency: examining physicians' perspectives toward online consumer reviews in the United States. *Int J Healthc Manag* 2020 Feb 20;14(4):1050-1056. [doi: [10.1080/20479700.2020.1728925](https://doi.org/10.1080/20479700.2020.1728925)]
37. Zhao HH, Luu M, Spiegel B, Daskivich TJ. Correlation of online physician rating subscores and association with overall satisfaction: observational study of 212,933 providers. *J Med Internet Res* 2020 Oct 27;22(10):e11258-e11258 [FREE Full text] [doi: [10.2196/11258](https://doi.org/10.2196/11258)] [Medline: [33107826](https://pubmed.ncbi.nlm.nih.gov/33107826/)]
38. Feel better about finding healthcare. Healthgrades. Healthgrades Marketplace URL: <https://www.healthgrades.com/find-a-doctor> [accessed 2021-07-29]

Abbreviations

- ADOC:** Alabama State Department of Corrections
BOP: US Bureau of Prisons
CDC: Centers for Disease Control and Prevention
PPE: personal protective equipment

Edited by G Eysenbach; submitted 03.08.21; peer-reviewed by M Behzadifar; comments to author 01.10.21; revised version received 04.10.21; accepted 04.10.21; published 10.11.21.

Please cite as:

Valera P, Carmona D, Malarkey S, Sinangil N, Owens M, Lefebvre A

Exploring Online Health Reviews to Monitor COVID-19 Public Health Responses in Alabama State Department of Corrections: Case Example

JMIR Form Res 2021;5(11):e32591

URL: <https://formative.jmir.org/2021/11/e32591>

doi: [10.2196/32591](https://doi.org/10.2196/32591)

PMID: [34609313](https://pubmed.ncbi.nlm.nih.gov/34609313/)

©Pamela Valera, David Carmona, Sarah Malarkey, Noah Sinangil, Madelyn Owens, Asia Lefebvre. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.11.2021. 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

Correction: The Potential of Digital Data Collection Tools for Long-lasting Insecticide-Treated Net Mass Campaigns in Nigeria: Formative Study

Youngji Jo¹, PhD; Nathan Barthel², MSc; Elizabeth Stierman², MPH; Kathryn Clifton², PhD; Esther Semee Pak³, MPH; Sonachi Ezeiru⁴; Diwe Ekweremadu⁴; Nnaemeka Onugu⁴; Zainab Ali⁴; Elijah Egwu⁴; Ochayi Akoh⁴, BSc; Orkan Uzunyayla⁵, MSc; Suzanne Van Hulle², MHS

¹Boston Medical Center, Boston, MA, United States

²Catholic Relief Services, Baltimore, MD, United States

³Graduate Institute of International Development Studies, Geneva, Switzerland

⁴Catholic Relief Services, Abuja, Nigeria

⁵Red Rose Ödeme Sistemleri AŞ, İstanbul, Turkey

Corresponding Author:

Youngji Jo, PhD

Boston Medical Center

One Boston Medical Center Pl

Boston, MA, 02118

United States

Phone: 1 4438001626

Email: youngji1435@gmail.com

Related Article:

Correction of: <https://formative.jmir.org/2021/10/e23648>

(JMIR Form Res 2021;5(11):e34526) doi:[10.2196/34526](https://doi.org/10.2196/34526)

In “The Potential of Digital Data Collection Tools for Long-lasting Insecticide-Treated Net Mass Campaigns in Nigeria: Formative Study” (JMIR Form Res 2021;5(10):e23648) the authors noted one error.

In the originally published manuscript, the name of author Suzanne Van Hulle was formatted incorrectly as the given names “Suzanne Van” and surname “Hulle.” This has now been corrected to the given name “Suzanne” and surname “Van

Hulle.” The citation information for this author has also been corrected from “Hulle SV” to “Van Hulle S.”

The correction will appear in the online version of the paper on the JMIR Publications website on November 8, 2021, 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 27.10.21; this is a non-peer-reviewed article; accepted 27.10.21; published 08.11.21.

Please cite as:

Jo Y, Barthel N, Stierman E, Clifton K, Pak ES, Ezeiru S, Ekweremadu D, Onugu N, Ali Z, Egwu E, Akoh O, Uzunyayla O, Van Hulle S

Correction: The Potential of Digital Data Collection Tools for Long-lasting Insecticide-Treated Net Mass Campaigns in Nigeria: Formative Study

JMIR Form Res 2021;5(11):e34526

URL: <https://formative.jmir.org/2021/11/e34526>

doi: [10.2196/34526](https://doi.org/10.2196/34526)

PMID: [34748514](https://pubmed.ncbi.nlm.nih.gov/34748514/)

©Youngji Jo, Nathan Barthel, Elizabeth Stierman, Kathryn Clifton, Esther Semee Pak, Sonachi Ezeiru, Diwe Ekweremadu, Nnaemeka Onugu, Zainab Ali, Elijah Egwu, Ochayi Akoh, Orkan Uzunyayla, Suzanne Van Hulle. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 08.11.2021. 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

Correction: Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland

Kristjana Thorarinsdottir¹, MSc; Emily A Holmes^{2,3}, Prof Dr; Johann Hardarson¹, MSc; Unnur Hedinsdottir¹, MSc; Marie Kanstrup³, PhD; Laura Singh^{2,4}, PhD; Arna Hauksdottir⁵, Prof Dr; Thorhildur Halldorsdottir^{1,5}, PhD; Berglind Gudmundsdottir^{6,7}, Prof Dr; Unnur Valdimarsdottir^{5,8,9}, Prof Dr; Edda Bjork Thordardottir⁵, PhD; Beau Gamble², PhD; Andri Bjornsson¹, Prof Dr

¹Department of Psychology, University of Iceland, Reykjavik, Iceland

²Department of Psychology, Uppsala University, Uppsala, Sweden

³Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

⁴Swedish Collegium for Advanced Study, Uppsala, Sweden

⁵Centre of Public Health Sciences, University of Iceland, Reykjavik, Iceland

⁶Faculty of Medicine, University of Iceland, Reykjavik, Iceland

⁷Landspítali - The National University Hospital of Iceland, Reykjavik, Iceland

⁸Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

⁹Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States

Corresponding Author:

Kristjana Thorarinsdottir, MSc

Department of Psychology

University of Iceland

Sæmundargata 12

Reykjavik, 101

Iceland

Phone: 354 6942001

Fax: 354 6942001

Email: kth35@hi.is

Related Article:

Correction of: <https://formative.jmir.org/2021/11/e29873>

(*JMIR Form Res* 2021;5(11):e34897) doi:[10.2196/34897](https://doi.org/10.2196/34897)

In “Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland” (*JMIR Form Res* 2021;5(11):e29873) the authors noted eight errors.

1. In the originally published paper, Edda Bjork Thordardottir's degree information was listed incorrectly as:

Edda Bjork Thordardottir, Prof Dr

This has been corrected to:

Edda Bjork Thordardottir, PhD

2. In the section “Participants,” the following sentence was included in the originally published paper:

The screening included a short description of the symptoms, followed by questions about the presence of the symptoms to assess their eligibility for this study.

This has been corrected to:

The screening included a short description of the symptom, followed by questions about the presence of the symptom to assess their eligibility for this study.

3. In the originally published paper, the final sentence of the section “Data Analysis” contained the following formula:

...this was calculated as $(1 - [12.6/6.1]) \times 100 = 52\%$ reduction...

This formula has been corrected to:

...this was calculated as $(1 - [6.1/12.6]) \times 100 = 52\%$ reduction...

4. In the section “Open Science Statement,” the following sentence was included in the originally published paper:

Study materials may be made available upon reasonable request with an appropriate materials transfer agreement with Uppsala University.

This has been corrected to:

Study materials may be made available upon reasonable request with an appropriate materials transfer agreement with University of Iceland.

5. In the section “Self-report Measures on PTSD, Depression and Anxiety Symptoms, and General Functioning,” the following sentence was included in the originally published paper:

Initial high levels of PTSD symptoms (a PCL-5 score of 51) were reduced by over half at postintervention, and the reduction was clearly clinically significant at the 3-month follow-up, with a score of only 5.

This has been corrected to:

Initial high levels of PTSD symptoms (a PCL-5 score of 51) were reduced by over half at postintervention, and the reduction was clearly clinically significant at the 3-month follow-up, with a score of only 6.

6. The caption of Figure 2 originally included the following sentence:

Gaps in the time series in the baseline and intervention periods reflect the missing data (for official intrusive memory data, see Figure 3).

This has been corrected to:

Gaps in the time series in the baseline and intervention periods reflect the missing data (for each specific intrusive memory data, see Figure 3).

7. In Table 3, footnote ^k originally read :

Have the intrusive memories affect your ability to function in your daily life?

This has been corrected to:

Have the intrusive memories affected your ability to function in your daily life?

8. In the originally published manuscript, the following text appeared in the Conflicts of Interest section:

EAH received funding from the Oak Foundation (OCAY-18-442) and from the Swedish Research Council (2020-00873) in support of this study; EAH also received funding from AFA Insurance (200342) and the Lupina Foundation. AB received funding from the Icelandic Research Fund (11709-0270). UV received funding to establish the stress and gene analysis (SAGA) cohort from the European Research Council (StressGene, grant 726413) and the Icelandic Research Fund (grant 163362-051). EBT reports funding from the Icelandic Research Fund (185287-051). LS received funding from the Swiss National Science Foundation (P2BEP1_184378) and a Thunberg Fellowship from the Swedish Collegium for Advanced Study.

This text has instead been moved to the Acknowledgments section of the paper, as it acknowledges funding sources.

The correction will appear in the online version of the paper on the JMIR Publications website on November 26, 2021, 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 11.11.21; this is a non-peer-reviewed article; accepted 11.11.21; published 26.11.21.

Please cite as:

Thorarinsdottir K, Holmes EA, Hardarson J, Hedinsdottir U, Kanstrup M, Singh L, Hauksdottir A, Halldorsdottir T, Gudmundsdottir B, Valdimarsdottir U, Thordardottir EB, Gamble B, Bjornsson A

Correction: Reducing Intrusive Memories of Childhood Trauma Using a Visuospatial Intervention: Case Study in Iceland
JMIR Form Res 2021;5(11):e34897

URL: <https://formative.jmir.org/2021/11/e34897>

doi: [10.2196/34897](https://doi.org/10.2196/34897)

PMID: [34889751](https://pubmed.ncbi.nlm.nih.gov/34889751/)

©Kristjana Thorarinsdottir, Emily A Holmes, Johann Hardarson, Unnur Hedinsdottir, Marie Kanstrup, Laura Singh, Arna Hauksdottir, Thorhildur Halldorsdottir, Berglind Gudmundsdottir, Unnur Valdimarsdottir, Edda Bjork Thordardottir, Beau Gamble, Andri Bjornsson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.11.2021. 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/>