JMIR Mental Health

Journal Impact Factor (JIF) (2022): 5.2
Volume 8 (2021), Issue 4 ISSN 2368-7959 Editor in Chief: John Torous, MD, MBI

Contents

Original Papers

Reviews

Original Paper

Implementing a Peer Advocate Mental Health Digital Intervention Program for Ohio Youth: Descriptive Pilot Study

Tashuna Albritton[1], MSW, PhD; Kelsey Lynett Ford[2], DrPH; Kira Elsbernd[2], MPH; Melodie Santodomingo[2], MPH; Ivan Juzang[3], MBA; Pam Weddington[3], BS; Sheana Bull[2], MPH, PhD

[1]CUNY School of Medicine, City College of New York, New York, NY, United States

[2]Colorado School of Public Health, University of Colorado, Aurora, CO, United States

[3]MEE Productions Inc, Philadelphia, PA, United States

Corresponding Author:
Tashuna Albritton, MSW, PhD
CUNY School of Medicine
City College of New York
160 Convent Ave
New York, NY, 10031
United States
Phone: 1 212 650 8224
Email: talbritton@med.cuny.edu

Abstract

Background: In the United States, millions of adolescents report poor mental health, where 1 in 5 teenagers considers suicide. Reducing stigma and fostering peer support remains critical for positive mental health interventions and programs. Increasingly, digital mental health tools have emerged with great promise, leveraging social networks. Despite the potential, limited understanding of such comprehensive programs and their implementation exist.

Objective: The objective of this study investigates a piloted digital mental health training program (Be Present) for youth, specifically describing the impact on youth behavioral outcomes and user engagement and identifying high-risk youth in the early phases of prevention.

Methods: Eligibility included Ohio residents (aged 14 to 22 years) to be enrolled as either a Friend or a Peer Advocate. From May 1 to June 1, 2019, participants completed the Advocate training course, taking pretest and posttest surveys. Single-arm descriptive analyses measured youth outcomes (self-efficacy, intentions, behaviors, social support, knowledge, and sources of strength) and engagement and assessed risk based on survey responses.

Results: A total of 65 adolescents participated, with 54 completing both pretest and posttest surveys. The majority of participants included non-Hispanic White females. Findings illustrated a significant increase in self-report of referrals for mental health services as well as in perceptions that youth had of experiencing social support; however, no significant differences were found for measures of self-efficacy, knowledge, and sources of strength between pretest and posttest surveys. Roughly two-thirds of the participants completed all of the Advocate training modules, and we observed a gradual decline in engagement. Most respondents who received escalated high-risk response messages identified as female.

Conclusions: The pilot presented promise for implementing a digital mental health program focused on peer support, specifically observing reported youth behavioral outcomes and user engagement and identifying high-risk youth. Various limitations exist given the small nonrepresentative sample and lack of control group. All findings should be considered preliminary to a larger trial and underscore the feasibility of delivering online training programs to bolster adolescent mental health. Such formative evaluation proved critical for future implementation and research, offering opportunity for substantial improvements for real-world digital mental health programs.

(JMIR Ment Health 2021;8(4):e24605) doi:10.2196/24605

KEYWORDS

mental health; adolescent; digital health; suicide prevention; social support; youth

Introduction

Background

In the United States every year, 1 in 5 teenagers considers suicide, and approximately 1 million teenagers have reported attempting suicide [1]. Suicide is the third leading cause of death for those aged 10 to 19 years [2]. People who considered suicide later reported the teenage years as the initial onset of suicidal ideation, making this time critical for a positive intervention [3]. Ohio's youth, defined as those aged 15 to 24 years, had a suicide rate of 11.27 per 100,000, which is comparable to the United States average for youth of the same age, at 11.39 per 100,000 [4]. Addressing mental illness and stigma associated with mental illness and equipping teens with skills to handle stressors associated with this period of life is critical for an intervention that targets suicide prevention.

Peer support is critical for any intervention targeting teenagers; this is especially true in mental health interventions. Youth who are suicidal are more likely to talk to other youth about being suicidal than to adults [5]. Understanding this peer support in teens can be coupled with mental health education to both raise awareness and challenge existing stigma associated with mental illness [6]. Social support has shown direct effects on mental health outcomes, particularly with peers [7].

Online social media networks may then increase the perceived social support along with increasing the information spread of user-generated content [8]. These online social networks can also be used to reach large numbers of teens as a platform for prevention and education interventions; 83% of young adults have reported using at least one social media site [9]. Frequent use of the internet, for teens, has been reported to increase communications both with family and other social support, beyond social media into face-to-face relationships [10]. Despite the potential of social networks and ability to support mental health initiatives with technology, limited understanding exists on how to implement these initiatives in the real world [11]. The objective of this study investigates a pilot digital mental health training program for youth, specifically describing impact on youth behavioral outcomes and user engagement and identifying high-risk youth in the early phases of prevention. Given the importance of this prevention program, it was equally important to conduct a program evaluation to assess the impact of the content on youth outcomes, its effectiveness in maintaining youth interest as a novel digital suicide prevention program, and its effect on youth most vulnerable to suicide.

Context

In response to ongoing concerns related to youth mental health, the Ohio Department of Mental Health and Addiction Services (Ohio MHAS) and the Motivational Educational Entertainment Productions Inc (MEE) team created Be Present, a digital e-learning platform designed to recruit and train youth to become peer mental health educators by facilitating protective factors to prevent mental health disorders that may lead to youth suicide. The e-learning program openly addresses identified stressors and traumas that put young people at risk for depression, bullying, peer pressure, substance use, and suicide. This is a report on the experiences of a group of participants who pilot-tested the online digital campaign. Participants in this pilot study completed a baseline survey, took part in a self-guided online e-learning course, and were surveyed again postintervention.

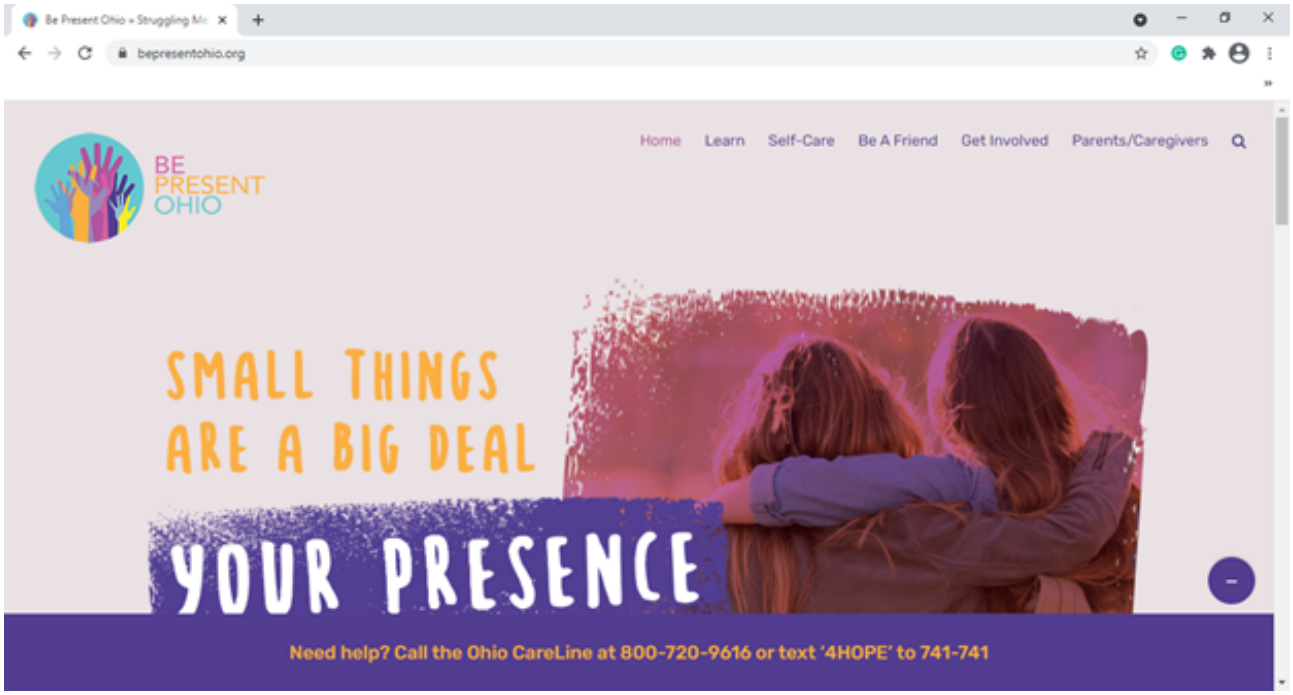
Methods

Be Present Campaign

In October 2017, MEE contracted with the mHealth Impact Lab of the Colorado School of Public Health to evaluate the impact of the e-learning campaign, Be Present, targeting youth suicide prevention in Ohio. Figure 1 displays the home page for the Be Present campaign. MEE trained an inaugural pilot cadre of Be Present Campaign Advocates as peer leaders. mHealth was tasked with evaluating youth (user) engagement with and impacts of the digital e-learning platform. They accomplished this through an assessment of change in self-reported mental health knowledge and protective factors assessed at the start and conclusion of the training.

Here we provide a description of this pilot assessment. This pilot study is classified as a program evaluation (Colorado Multiple Institutional Review Board); therefore, it was not filtered through formal institutional review board processes, although MEE and the mHealth Impact Lab adhered to ethical human subject practices in the execution of the pilot program and evaluation.

Figure 1. Be Present advocate training online platform.



Eligibility Criteria

The Be Present e-learning platform has options for 2 levels of participation open to any Ohio resident aged 14 to 22 years. The first level is called a Be Present Friend, open to those completing the online registration available on the campaign website [12]. After becoming a Friend, youth could participate in the second level and become trainees for Be Present Advocates (ie, peer educators who facilitate youth mental health in their social networks). These Advocates were eligible to complete the comprehensive pretraining and posttraining assessments of this pilot program. The registration process for the Advocate trainees included completing an informed consent or assent form accompanied by parental approval (for youth aged younger than 18 years) and identifying a nonparental adult (meaning not their parent) advisor to support the training process.

Recruitment

Participants were recruited by MEE primarily through school and community events. Additional recruitment occurred via social and digital media marketing that highlighted the availability of Advocate training. Social media sites such as Facebook, Instagram, Snapchat, and Twitter along with the Be Present website and the campaign's dedicated YouTube channel were used to recruit youth.

Data Collection

All data were collected from May 1 to June 1, 2019, by MEE, anonymized, and securely transmitted to mHealth for comprehensive analysis. Several types of data were collected from the Be Present Campaign Advocate training course: (1) data from the pretraining and posttraining surveys (answers to multiple-choice questions) and (2) selected Advocate trainee–entered information in response to questions and assignments in the lessons and modules (answers to multiple-choice and open-ended questions, uploads of assignments, watching instructional videos, etc).

Youth Advocate trainees completed an enrollment survey on the e-learning site before starting the course. After successful enrollment in the Advocate training program, youth watched educational videos and completed online lessons and modules at their own pace. Data on youth engagement with the training curriculum (eg, watching videos, using the safety e-toolbox, completing homework assignments, and sharing messages on social media) were gathered on the training website and maintained by MEE. The social media sites and handles that youth provided during registration were monitored by MEE to determine that assignments requiring a social media share were being completed.

When they were close to completing the online content, Advocate trainees were invited to attend a single community-based training session offered multiple times between May 1 and June 1, 2019, for participant convenience. Attendance was mandatory to receive full Advocate certification. Once Advocate trainees completed the in-person training and while still on-site, they were required to complete a second online survey (posttest) with the same information gathered at baseline. Data from the survey were maintained by mHealth. Following completion of the second survey, trainees were fully certified as Be Present Campaign Advocates.

The pretest and posttest survey, taken anonymously, was developed using Qualtrics. Validated survey questions (Multimedia Appendix 1) were revised from the Rosenberg Self-Esteem Scale [13] and the Community Attitudes Toward Mental Illness questionnaire [14] for measuring mental health stigma, sources of strength, and social support. Additional questions captured standard demographics information, youth resilience measures, positive and negative coping, self-efficacy, and ways to be present for family and friends.

Measures

Self-report surveys were used to measure 7 youth outcomes: self-efficacy, intentions, behaviors, social support, knowledge, self-esteem, and sources of strength. See Table 1 for a brief description of each youth outcome.

Scores were calculated for each survey domain: self-efficacy, intentions, behaviors, social support, knowledge, self-esteem, and sources of strength. Scores for each survey domain were created by summing scores for each domain question (0=best response, 1=good response, 2=poor response, 3=poorest response). Lower scores indicate better responses.

Table 1. Description of youth self-report survey measures.

Outcomes	Description
Self-efficacy	Proportion of respondents who report high measures of their ability to succeed in different situations and tasks. Scored responses report belief and capacity to execute behaviors.
Intentions	Proportion of respondents who report high measures of planned intentions around goals. Scored responses report the aim or plan to execute behaviors.
Behaviors	Proportion of respondents who report referrals to mental health resources for themselves or peers.
Social support	Proportion of respondents who report high measures of social support from family, peers, and community. Scored responses report supportive individuals and community.
Knowledge	Proportion of respondents who report high measures of knowledge around self-care and mental health. Scored responses report knowledge of mental health topics.
Self-esteem	Proportion of respondents with self-satisfaction rated at 3 or more and assessments of feeling no good, useless, or a failure at 2 or fewer on a 5-point scale.
Sources of strength	Proportion of respondents who report high measures of peer social networks. Scored responses report peer social networks.

Engagement

MEE also captured information from the website portal to the training course. During the self-guided online training for each participant, MEE collected user data that reflects engagement with the training content (ie, lessons and modules, completion of modules) and how that information is internalized by trainees and then shared by them with their adult advisors and peers, either in person or via their existing social media networks. MEE also collected additional data from the Advocate training course, including demographic information on users. A descriptive summary of enrollment, demographic information, and overall user engagement was generated for each completed training module.

High Risk

We assessed whether youth were at risk themselves for a negative mental health outcome as they responded to questions. These risk-assessing questions were in the behaviors, social support, self-esteem, and sources of strength domains. Participants received an escalated high-risk response message during the survey if they selected disagree or strongly disagree on answers outlined in Multimedia Appendix 1. Each domain had separate criteria; participant had to disagree with all statements within a domain to receive a high-risk message. The high-risk response message was worded as follows:

> *We hear you. It's OK to not be OK. We thank you for your openness and honesty in answering these survey questions. Everyone has a bad day from time to time. But sometimes, something more might be going on. If you're feeling hopeless, overwhelmed, or in a crisis, it might be time to get some immediate support to help you get through it. When the emotional pain seems too big to handle, get help. Start with these resources where people will step up and Be Present for you... To talk to someone, text 4HOPE to the Crisis Text Line (741741) or call the National Suicide Prevention Lifeline (1-800-273-8255). Both are available 24/7.*

The algorithm and high-risk response messages were vetted by clinical and research experts. Each respondent could receive up to 4 escalated high-risk response messages per domain.

Analysis

All statistical analysis was performed using SAS 9.4 (SAS Institute Inc). The study data used in this analysis were collected by MEE team members and managed using Qualtrics survey software hosted at the University of Colorado Anschutz Medical Campus, where mHealth resides. All data were anonymized and reviewed for completeness and consistency. Descriptive statistics were performed to understand the study population, summarize enrollment and engagement, and detect any erroneous values.

A descriptive assessment was performed integrating Be Present e-learning engagement data from MEE with single-arm, pretest, and posttest assessments with Be Present Advocates. The Student *t* test was used to assess differences in mean scores on each of the survey domains between pretest and posttest. $P<.05$ was considered to be statistically significant. Engagement with the digital e-learning platform was measured as the percentage of participants who completed each of the training blocks. Changes in self-reported intentions, self-efficacy, and behaviors related to mental health were measured at baseline and postparticipation in the Be Present Campaign in a single-arm design. The study enrolled 65 participants, of which 62 had data from the baseline survey and 54 had data from both baseline and follow-up surveys. Only participants who had baseline data were included in the analysis.

High-risk responses were also analyzed using descriptive analytics to determine the characteristics of respondents who received escalated high-risk messages during the survey completion. These were survey questions identifying high-risk respondents. High-risk messages were positioned in the domains behaviors, social support, self-esteem, and sources of strength.

Results

Demographics

Table 2 reports the characteristics of Be Present Campaign Advocate trainees who completed the pretraining (baseline) survey. Of the respondents, 82% (51/62) identified as female and 95% (59/62) as straight or heterosexual. A total of 77% (48/62) of respondents were non-Hispanic White, 10% (6/62) identified as Hispanic, and 13% (8/62) as non-Hispanic non-White.

Table 2. Demographic characteristics of participants in the Be Present online training intervention pilot test for youth suicide prevention at baseline, May-June 2019 (n=62)[a].

Characteristic	Value, n (%)
Gender	
Male	11 (18)
Female	51 (82)
Other	—[b]
Prefer not to answer	—
Sexual orientation	
Straight or heterosexual	59 (95)
Lesbian, gay, or homosexual	1 (2)
Bisexual	1 (2)
Prefer not to answer	1 (2)
Race/ethnicity	
Hispanic or Latino	6 (10)
Non-Hispanic White	48 (77)
Non-Hispanic non-White	8 (13)

[a]Percentages may not add up to 100 due to rounding.

[b]Not applicable.

Engagement

This is a descriptive summary of user engagement in the Be Present online training pilot test for youth suicide prevention, May-June 2019. Of the participants enrolled, 60% (37/62) completed all 7 training blocks.

The completion percentages for each training block were as follows: self-efficacy (54/62, 87%), intentions (53/62, 85%), behaviors (51/62, 82%), social support (45/62, 73%), knowledge (37/62, 60%), self-esteem (37/62, 60%), and sources of strength (37/62, 60%).

Youth Behavioral Outcomes

Table 3 illustrates the percentage improvement in scores for each survey domain. There were statistically significant improvements in behaviors (2.75, P=.007) and social support (4.13, P<.001). All other categories saw modest improvement from pretest to posttest.

Table 3. Changes in domain scores between baseline and postintervention surveys.

Domain	Pretest, mean	Posttest, mean	Percentage change in score	*P* value
Self-efficacy	13.50	14.31	0.81	.42
Intentions	7.38	7.60	0.22	.83
Behaviors	31.85	34.60	2.75	.007
Social support	62.67	66.80	4.13	.001
Knowledge	24.98	25.98	1.00	.32
Self-esteem	21.48	21.78	0.30	.76
Sources of strength	23.04	24.93	1.89	.06

High-Risk Responses

Table 4 outlines the proportion of participants who selected answers that prompted an escalated high-risk response message. Of the respondents who completed the pretest, 15% (9/62) of respondents received a total of 13 escalated high-risk response messages. Five respondents received an escalated high-risk response message, triggered when they indicated engaging in risky behaviors or having limited social support, self-esteem, or sources of strength. Each respondent could receive up to 4 escalated high-risk messages. The remaining survey domains do not have questions identifying high-risk respondents. The most escalated response messages were to females across the survey domains social support (n=5), self-esteem (n=6), and sources of strength (n=1). Among respondents who completed both the pretest and posttest survey, 9% (5/54) received a total of 7 escalated high-risk messages in the posttest survey. Females received the most escalated response messages across the survey domains: behaviors (n=2) and self-esteem (n=4).

Table 4. High-risk responses by category, pretest and posttest.

Domain	Pretest		Posttest		Total
	Male, n (%)	Female, n (%)	Male, n (%)	Female, n (%)	
Behaviors	—[a]	—	—	2 (33)	2
Social support	1 (1)	5 (42)	—	—	6
Self-esteem	—	6 (50)	1 (1)	4 (67)	11
Sources of strength	—	1 (8)	—	—	1
Total	1 (5)	12 (60)	1 (5)	6 (30)	20

[a]Not applicable.

Discussion

Principal Findings

The objective of this study was to present descriptive results of a primarily digital pilot mental health training program for Ohio youth. For this paper, we focused specifically on the impact on youth behavioral outcomes, user engagement, and identifying those at high-risk for immediate intervention. Our results are process-focused by design as a way of providing information to improve the overall program and implementation.

Youth Behavioral Outcomes

Youth behavioral outcomes were assessed immediately posttraining, which reflects an immediate short-term outcome rather than a sustained immediate or longer term outcome. We did observe a significant increase in self-report of referrals for mental health services as well as in perceptions that youth had of experiencing social support. These results offer some initial optimism considering that the Be Present program may have potential for impact, although we cannot suggest this is a definitive conclusion without a larger and more rigorous examination of the program.

In the other domains explored, including intentions, self-efficacy, knowledge, and sources of strength, we observed no significant difference between pretest and posttest self-assessments for youth. Given that we have a small sample and no control group, our findings should be considered only preliminary to a larger trial and underscore primarily the feasibility of delivering online training programs to bolster youth mental health.

Engagement

Close to two-thirds of the enrollees completed all of the Advocate training modules. The first 3 training blocks (blocks 1, 2, and 3) had a higher percentage of enrollee completion compared with the latter 3 blocks (blocks 4, 5, and 6). A systematic review of mobile and web-based interventions indicated that engagement decreases in interventions that have a longer duration [15]. Although we are uncertain as to why there was a gradual decline in engagement, we speculate that intervention duration may have played a role. Also, enrollees were allowed to complete the modules at their own pace, which may have negatively impacted sustained interest and engagement. A solution might include reducing the number of training modules and setting a time limit for module completion.

High-Risk Responses

Most respondents who received escalated high-risk response messages identified as female. Due to the small number of male respondents, this may not be representative of the larger population. However, evidence has shown that adolescent and adult females report a higher percentage of suicidal ideations compared with their male counterparts [16,17]. A full-scale intervention evaluation examining high-risk responses is required to suggest gender differences among Ohio adolescents.

Limitations and Strengths

Although we discovered interesting descriptive findings that will inform modifications to the next iteration of Be Present, the overall study is not without limitations. The majority of respondents in this pilot identified as white, female, and straight or heterosexual. While reflective of the larger school population, this is not generalizable. Only 60% of enrollees completed the whole training with no indication as to why other enrollees did not complete the training. This introduces the potential for biased results as we do not know the impetus for or against completing the training.

Self-report surveys yielded questionable responses as respondents answered every question identically in several cases, raising concern with the study's external validity. This is a limitation of the pretest and posttest evaluation. However,

exploration of ways to improve the accuracy of self-report among adolescents may be helpful.

We also recognize process-related limitations. The survey should have been streamlined to include one link to improve the delivery of questions, reduce data processing requirements, and avoid distribution issues. Not enough time was allotted between the training and surveys to draw more conclusive results on the training's impact. Although we can describe engagement with Advocate training data, we cannot make conclusions on effectiveness. Future steps should include the possibility of evaluating effectiveness and efficacy, similar to other digital mental health pilot programs [18].

One of several strengths of the study is that it provides insights into future implementation for the Be Present program. These data will be used to modify the training program and implementation to improve both process and engagement and will be useful for understanding the impact of mental health advocacy at certain points in youth development. Finally, this study has implications for the future of youth mental wellness social media campaigns and youth engagement.

Future Work

We plan to implement the Be Present program in other areas using the insights gained from this pilot. Next steps in the Be Present program include the addition of more diverse populations to evaluate behavioral outcomes and engagement. Furthermore, a more streamlined process will allow for more time between Advocate training and outcome measures. Our goal is to test for the preliminary efficacy of the program. Follow-up studies from this pilot will be robust enough to allow for more complex statistical approaches to truly capture changes in self-report attitudes toward mental health and determine if engagement with the campaign is associated with youth behavioral outcomes. Although this analysis cannot be used to make full conclusions on the impact on youth behavioral outcomes, it does give key insight into the impact of digital training platforms.

Conclusion

In this paper, we provided descriptive analyses for a peer advocate mental health digital intervention program for youth in Ohio. The important take-away from this study is that it takes time to develop a solid digital mental health program, especially for adolescents. Those who venture into developing prevention programs must be prepared for more than one iteration of a pilot. Factors beyond what is discussed in this article confound the process for a seamless implementation—duration uncertainties, relatability to the content, range in developmental maturity, access to the internet and technology devices, and other competing factors like schoolwork and family obligations. Although the internet and ready availability of devices make programs more accessible, adolescents are typically still at the mercy of a parent who can determine if and when their child can use the internet and devices. For this reason, it is beneficial for prevention intervention programs to include a formative process evaluation in which one can monitor engagement and implementation and make the necessary changes as they go. Program development is time-consuming and expensive, and thus modifying as one goes is prudent.

Acknowledgments

We would like to acknowledge MEE Productions Inc and extend a special thank you to the pilot high school and its principal.

Conflicts of Interest

None declared.

Multimedia Appendix 1
High-risk assessment.
[DOCX File , 19 KB - mental_v8i4e24605_app1.docx]

References

1. Grunbaum J, Kann L, Kinchen S, Williams B, Ross J, Lowry R, et al. Youth risk behavior surveillance—United States, 2001. J Sch Health 2002 Oct;72(8):313-328. [doi: 10.1111/j.1746-1561.2002.tb07917.x] [Medline: 12389372]
2. Anderson RN. Deaths: leading causes for 2000. Natl Vital Stat Rep 2002 Sep 16;50(16):85 [FREE Full text] [Medline: 12355905]
3. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. Arch Gen Psychiatry 1999 Jul;56(7):617-626. [doi: 10.1001/archpsyc.56.7.617] [Medline: 10401507]
4. Youth Suicide in Ohio. Violence and Injury Prevention Program, Ohio Department of Health. 2016 Oct. URL: https://odh.ohio.gov/wps/wcm/connect/gov/7f853eb4-9c94-4caf-9dc8-208d3bb92a7c/FS-Youth-Suicide-in-Ohio-Final.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=ROOTWORKSPACE.Z18_M1HGGIK0N0JO00QO9DDDDM3000-7f853eb4-9c94-4caf-9dc8-208d3bb92a7c-mkltZPB [accessed 2021-04-05]
5. Gould G, Greenberg T, Velting DM, Shaffer D. Youth suicide risk and preventive interventions: a review of the past 10 years. J Am Acad Child Adolesc Psychiatry 2003 Apr;42(4):386-405. [doi: 10.1097/01.CHI.0000046821.95464.CF] [Medline: 12649626]

6. Essler V, Arthur A, Stickley T. Using a school-based intervention to challenge stigmatizing attitudes and promote mental health in teenagers. J Mental Health 2009 Aug 26;15(2):243-250. [doi: 10.1080/09638230600608669]
7. Colarossi L, Eccles J. Differential effects of support providers on adolescents' mental health. Soc Work Res 2003 Mar 01;27(1):19-30. [doi: 10.1093/swr/27.1.19]
8. Chou W, Hunt Y, Beckjord E, Moser R, Hesse B. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Nov 27;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
9. Best P, Manktelow R, Taylor B. Online communication, social media and adolescent wellbeing: a systematic narrative review. Children Youth Serv Rev 2014 Jun;41:27-36. [doi: 10.1016/j.childyouth.2014.03.001]
10. Lee S. Online communication and adolescent social ties: who benefits more from internet use? J Comput Mediated Comm 2009;14(3):509-531. [doi: 10.1111/j.1083-6101.2009.01451.x]
11. Liverpool S, Mota C, Sales C, 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] [Medline: 32442160]
12. Be Present Ohio. URL: https://bepresentohio.org/ [accessed 2021-04-06]
13. Rosenberg M. Society and the Adolescent Self-Image. Princeton: Princeton University Press; 1965.
14. Taylor S, Dear M. Scaling community attitudes toward the mentally ill. Schizophr Bull 1981;7(2):225-240. [doi: 10.1093/schbul/7.2.225] [Medline: 7280561]
15. Nelson L, Coston T, Cherrington A, Osborn C. Patterns of user engagement with mobile- and web-delivered self-care interventions for adults with T2DM: a review of the literature. Curr Diab Rep 2016 Jul;16(7):66 [FREE Full text] [doi: 10.1007/s11892-016-0755-1] [Medline: 27255269]
16. Nowotny K, Peterson R, Boardman J. Gendered contexts: variation in suicidal ideation by female and male youth across U.S. states. J Health Soc Behav 2015 Mar;56(1):114-130 [FREE Full text] [doi: 10.1177/0022146514568350] [Medline: 25722128]
17. Burstein B, Agostino H, Greenfield B. Suicidal attempts and ideation among children and adolescents in US emergency departments, 2007-2015. JAMA Pediatr 2019 Jun 01;173(6):598-600 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.0464] [Medline: 30958529]
18. Lattie E, Ho J, Sargent E, Tomasino K, Smith J, Brown C, et al. Teens engaged in collaborative health: the feasibility and acceptability of an online skill-building intervention for adolescents at risk for depression. Internet Interv 2017 Jun;8:15-26 [FREE Full text] [doi: 10.1016/j.invent.2017.02.003] [Medline: 28584734]

Abbreviations

MEE: Motivational Educational Entertainment Productions Inc
MHAS: Ohio Department of Mental Health and Addiction Services

Edited by J Torous; submitted 26.09.20; peer-reviewed by S Liverpool, D Reidenberg; comments to author 01.11.20; revised version received 29.01.21; accepted 05.02.21; published 23.04.21.

Please cite as:
Albritton T, Ford KL, Elsbernd K, Santodomingo M, Juzang I, Weddington P, Bull S
Implementing a Peer Advocate Mental Health Digital Intervention Program for Ohio Youth: Descriptive Pilot Study
JMIR Ment Health 2021;8(4):e24605
URL: https://mental.jmir.org/2021/4/e24605
doi:10.2196/24605
PMID:33890868

Original Paper

A Smartphone-Based Self-management Intervention for Bipolar Disorder (LiveWell): User-Centered Development Approach

Geneva K Jonathan[1], MS; Cynthia A Dopke[1], PhD; Tania Michaels[2], MD; Andrew Bank[1], BA; Clair R Martin[1], BA; Krina Adhikari[1], MA, LPC; Rachel L Krakauer[3], MA; Chloe Ryan[4], LCSW; Alyssa McBride[1], BA; Pamela Babington[1], BA; Ella Frauenhofer[1], BA; Jamilah Silver[5], BA; Courtney Capra[1], BA; Melanie Simon[6], BA; Mark Begale[7], BA; David C Mohr[1], PhD; Evan H Goulding[1], MD, PhD

[1]Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]Pediatrics, Loma Linda Children's Hospital, Loma Linda, CA, United States

[3]Department of Psychology, University of Regina, Regina, SK, Canada

[4]Department of Social Work, UPMC Western Psychiatric Hospital, Pittsburgh, PA, United States

[5]Department of Psychology, Stony Brook University, Stony Brook, NY, United States

[6]Department of Psychology, School of Science and Engineering, Tulane University, New Orleans, LA, United States

[7]Vibrent Health, Fairfax, VA, United States

Corresponding Author:
Evan H Goulding, MD, PhD
Department of Psychiatry and Behavioral Sciences
Feinberg School of Medicine
Northwestern University
Chicago, IL
United States
Phone: 1 3125031189
Email: e-goulding@fsm.northwestern.edu

Abstract

Background: Bipolar disorder is a serious mental illness that results in significant morbidity and mortality. Pharmacotherapy is the primary treatment for bipolar disorder; however, adjunctive psychotherapy can help individuals use self-management strategies to improve outcomes. Yet access to this therapy is limited. Smartphones and other technologies have the potential to increase access to therapeutic strategies that enhance self-management while simultaneously providing real-time user feedback and provider alerts to augment care.

Objective: This paper describes the user-centered development of LiveWell, a smartphone-based self-management intervention for bipolar disorder, to contribute to and support the ongoing improvement and dissemination of technology-based mental health interventions.

Methods: Individuals with bipolar disorder first participated in a field trial of a simple smartphone app for self-monitoring of behavioral targets. To develop a complete technology-based intervention for bipolar disorder, this field trial was followed by design sessions, usability testing, and a pilot study of a smartphone-based self-management intervention for bipolar disorder. Throughout all phases of development, intervention revisions were made based on user feedback.

Results: The core of the LiveWell intervention consists of a daily self-monitoring tool, the Daily Check-in. This self-monitoring tool underwent multiple revisions during the user-centered development process. Daily Check-in mood and thought rating scales were collapsed into a single wellness rating scale to accommodate user development of personalized scale anchors. These anchors are meant to assist users in identifying early warning signs and symptoms of impending episodes to take action based on personalized plans. When users identified personal anchors for the wellness scale, the anchors most commonly reflected behavioral signs and symptoms (40%), followed by cognitive (25%), mood (15%), physical (10%), and motivational (7%) signs and symptoms. Changes to the Daily Check-in were also made to help users distinguish between getting adequate sleep and keeping a regular routine. At the end of the pilot study, users reported that the Daily Check-in made them more aware of early warning signs and symptoms and how much they were sleeping. Users also reported that they liked personalizing their anchors and plans and felt this process was useful. Users experienced some difficulties with developing, tracking, and achieving target goals. Users also did not consistently follow up with app recommendations to contact providers when Daily Check-in data suggested they needed

additional assistance. As a result, the human support roles for the technology were expanded beyond app use support to include support for self-management and clinical care communication. The development of these human support roles was aided by feedback on the technology's usability from the users and the coaches who provided the human support.

Conclusions: User input guided the development of intervention content, technology, and coaching support for LiveWell. Users valued the provision of monitoring tools and the ability to personalize plans for staying well, supporting the role of monitoring and personalization as important features of digital mental health technologies. Users also valued human support of the technology in the form of a coach, and user difficulties with aspects of self-management and care-provider communication led to an expansion of the coach's support roles. Obtaining feedback from both users and coaches played an important role in the development of both the LiveWell technology and human support. Attention to all stakeholders involved in the use of mental health technologies is essential for optimizing intervention development.

(JMIR Ment Health 2021;8(4):e20424) doi:10.2196/20424

KEYWORDS

behavioral intervention technology; mHealth; bipolar disorder; depression; illness management; smartphone; behavior change; early warning signs; self-management; qualitative

Introduction

Bipolar disorder is characterized by recurrent episodes of mania, hypomania, depression, and mixed states. Many individuals experience multiple acute episodes, long episode durations, and interepisode symptoms [1-5]. As a result, individuals with bipolar disorder are symptomatically ill about half of the time, and three-quarters of those affected never achieve full recovery of psychosocial function [1,6]. Pharmacological management is the primary treatment for bipolar disorder and effectively reduces relapse risk and symptom burden [7,8]. The addition of empirically supported psychotherapy to pharmacotherapy can further lower relapse rates, decrease symptoms, and improve quality of life [9-17].

Although psychotherapy delivered during research trials has proven helpful, only about half of individuals with bipolar disorder receive any therapy [18,19]. Enhancing access to content and tools from empirically supported psychotherapy is essential to improving treatment for bipolar disorder. Digital mental health technologies, including web and smartphone-based applications, provide a means to increase the availability of empirically supported psychotherapeutic strategies for managing mental health problems [20-24]. In addition to access barriers, current tools for bipolar disorder self-management can be improved [14,25-28]. Relative to face-to-face treatment, mental health technologies can advance intervention functionality by providing real-time assessments, feedback, and provider alerts [29]. Many people with bipolar disorder are interested in utilizing self-management strategies to stay well, and strategies used by individuals who are doing well overlap significantly with the content of empirically supported psychotherapies for bipolar disorder [30-33]. This overlap suggests that smartphone-based interventions delivering self-management strategies derived from empirically supported psychotherapies may meet user needs and support user engagement [34-37].

To address the need for increased access to and enhancement of empirically supported therapy for bipolar disorder, we designed and developed LiveWell, a smartphone-based self-management intervention (Clinicaltrials.gov NCT02405117, NCT03088462). Description of intervention development is essential to support ongoing improvement and dissemination of technology-based mental health interventions [38-42], and thus, this paper describes the user-centered development of LiveWell. The development approach aims to ground LiveWell in the lived experiences of individuals with bipolar disorder to create an intervention that encourages the development and long-term use of self-management strategies for living well with bipolar disorder [43-45]. The user-centered development was carried out in phases, consisting of (1) an initial 12-week field trial of a simple smartphone app for self-monitoring of behavioral targets, (2) design interviews and usability testing of a self-management app, and (3) an 8-week pilot trial of the complete LiveWell intervention, including both technology and human support.

Methods

Users

The study was reviewed and approved by Northwestern University's institutional review board. Users were recruited via fliers placed at university-affiliated and private outpatient mental health practices. Eligible users were 18-65 years old and had a DSM-IV (Diagnostic and Statistical Manual of Mental Disorders–fourth version) diagnosis of bipolar disorder with a minimum of 2 acute mood episodes within 2 years of enrollment. Individuals were excluded if they (1) were not in current psychiatric care; (2) met criteria for a substance-use disorder within the last 6 months; (3) met criteria for another psychiatric diagnosis, or had symptoms for which participation in the study was either inappropriate or dangerous (including current severe suicidal ideation or a serious suicide attempt in the last 12 months); (4) were pregnant or planned to become pregnant; (5) had a visual, hearing, voice, or motor impairment that would prevent completion of the study procedures or limit smartphone use; or (6) were unable to speak or read English. As the intervention primarily targets mood episode relapse prevention, a current mood episode at the baseline assessment was an additional exclusion criterion for the pilot study.

Before an initial telephone screening, users completed informed consent by telephone or online. Users completed written consent before a face-to-face (F2F) clinic visit. Initial screening to establish a bipolar disorder diagnosis was conducted via

telephone using the Mini–International Neuropsychiatric Interview [46]. If eligible, users completed an F2F interview with a study clinician (psychiatrist or psychologist) using an abbreviated version of the Affective Disorders Evaluation and the Clinical Monitoring Form [47,48]. Individuals with a confirmed diagnosis at the clinic visit enrolled in the field trial, design interview, and/or usability testing and could additionally continue to the next phase if they chose. For the pilot study, users also completed a baseline telephone assessment using the Clinical Monitoring Form to assess current clinical status. All individuals who participated in earlier phases of intervention development were offered an opportunity to participate in subsequent development phases. Users were compensated for their time and travel costs: US $10 was given toward travel costs and the telephone assessment; US $15 was given for participation in the clinical assessment, baseline and monthly telephone assessment, exit interview, and app training; and US $25 was given for the design and usability interviews.

Procedures

Field Trial

A 12-week field trial was completed by 4 users to assess a simple self-monitoring smartphone app. The research team collaborated with the Center for Behavioral Intervention Technologies (CBITS) at Northwestern University to develop an Android app with self-report data collection, encryption, and transmission to a secure server. Users were provided with a smartphone with a data plan and had an F2F meeting with a study staff member (coach) who used a structured script and handouts to introduce the app (Multimedia Appendix 1). The coach used structured interview scripts to gather feedback about the training after the session and app use after the field trial (Multimedia Appendix 1).

Design Sessions and Usability Testing

Design sessions were conducted with 4 users from the initial study and 1 additional user, for which a structured script and handouts (Multimedia Appendix 2) were used to simulate an F2F app training session with a coach. Users were provided with a smartphone app mock-up (Multimedia Appendix 3), asked to imagine using the app for the next 16 weeks, and instructed to think and ask any questions out loud. The research team collaborated with CBITS to extend the self-monitoring Android app to include information and tools to help users engage in bipolar disorder self-management. The 5 users from the design sessions attended F2F usability testing sessions that employed structured scripts and scenarios (Multimedia Appendix 4). Users were given 5-10 minutes to explore the app and asked to share their general impressions. Then, users read one assigned lesson and another of their choice and were asked to provide feedback about the lesson's usefulness, length, and coverage. Next, users read aloud scenarios that mimicked real-life situations: medication nonadherence, sleeping too little, and experiencing early warning signs of depression. After reading each scenario, users completed a daily check-in based on the scenario and received automated feedback from the app based on their self-report data. Users were asked to discuss the usefulness of this feedback and how it could be improved. All users were given a posttask questionnaire to rate the overall usability of the app (Multimedia Appendix 5).

Pilot Study

To test the complete intervention, 11 users, including 4 who attended the usability testing, completed an 8-week pilot study. Users were provided with a smartphone and a data plan. They had an F2F meeting with a coach who used a structured script and handouts (Multimedia Appendix 6) to instruct them on the use of the app and the coach's role. Following this meeting, users completed 6 phone calls (weeks 1-4, 6, and 8) with a coach; the coach used structured scripts (Multimedia Appendix 6) to support app use adherence, self-management strategy use (including the development of personalized wellness plans; Multimedia Appendix 7), and communication with clinical care providers. After completing the pilot, users completed a structured exit interview (Multimedia Appendix 8) and an exit questionnaire (Multimedia Appendix 9) to provide feedback about the app’s usability.

Analysis

Instant Data Analysis

The research team utilized an instant data analysis approach across all development phases to make iterative changes to the technology and coaching based on user feedback [49]. This approach reduces the time needed for analysis while also identifying usability issues [50,51]. Immediately after observing, interacting with, or receiving feedback from users, study staff wrote memos documenting users' problems or comments. The design team (coaches, programmers, the project manager, and team leaders) discussed these memos until they achieved consensus on necessary changes [52]. These discussions sometimes prompted a return to the literature to provide additional information to make design decisions. During the pilot study, coaches' experiences with the technology and the users were increasingly incorporated into the design discussions and decisions. Based on the number of users, the field trial and usability studies (n=4-5) should identify 55%-85% of problems with app usability, and the pilot study (n=11) should detect 80%-95% of usability problems [53-56]. In addition to the traditionally defined users, 4 coaches and 1 team leader provided human support during different stages of the development process.

Analysis of Usability Testing Posttask and Pilot Study Exit Questionnaires

Participant responses to the usability testing posttask questionnaire (n=5) and pilot study exit questionnaire (n=11) were used to assess usability (Multimedia Appendices 5 and 9). To provide summary assessments, responses from the 7-point response scales were collapsed into 2 categories: disagree/strongly disagree and agree/strongly agree.

Analysis of Pilot Study Exit Interviews

The exit interviews (n=11) were transcribed verbatim and used for thematic analysis [57]. Then, 3 researchers independently conducted a preliminary round of coding during which transcripts were partitioned into excerpts (transcript lines conveying a codable unit) and exported to Microsoft Excel

spreadsheets (Multimedia Appendix 10). App usability subthemes were inductively coded and deductively grouped into larger themes. Coders used nominal group consensus, meeting with a moderator to discuss, clarify differences in, and finalize codes [52].

Analysis of Rating Scale Anchors

Thematic analysis was used to investigate users' personalized anchors for the mood and thought rating scales used during the field trial (n=4) and the wellness rating scales used during the design sessions (n=5) and pilot study (n = 11) [57]. Anchors were entered into Excel spreadsheets (Multimedia Appendix 11), and 2 researchers inductively coded subtypes. A third researcher reviewed the codes, and a consensus process was used to finalize anchor subtypes [52]. Anchor subtypes were then deductively grouped into broader categories (anchor types) based on a literature review describing early warning signs experienced by individuals with bipolar disorder prior to an episode [11,12]. The research team discussed the overall coding scheme and developed definitions and examples of the anchor types and subtypes. Then, 2 researchers who were not involved in the initial coding and development used these definitions and examples to code the anchor types and subtypes across all 3 scales; the joint probability of agreement was 90% for subtypes and 87% for types.

Results

Users

At study enrollment, 12 users were included and were 21-62 (mean 38, SD 14) years old. Of these 12 users, 4 were male and 8 were female; 12 were non-Hispanic white; 3 were married or living as married, 3 were divorced, and 6 were never married; 5 had completed some college, 2 had completed college, and 5 had completed more than college; 2 were students, 6 were employed, 2 were unemployed, and 2 were on disability.

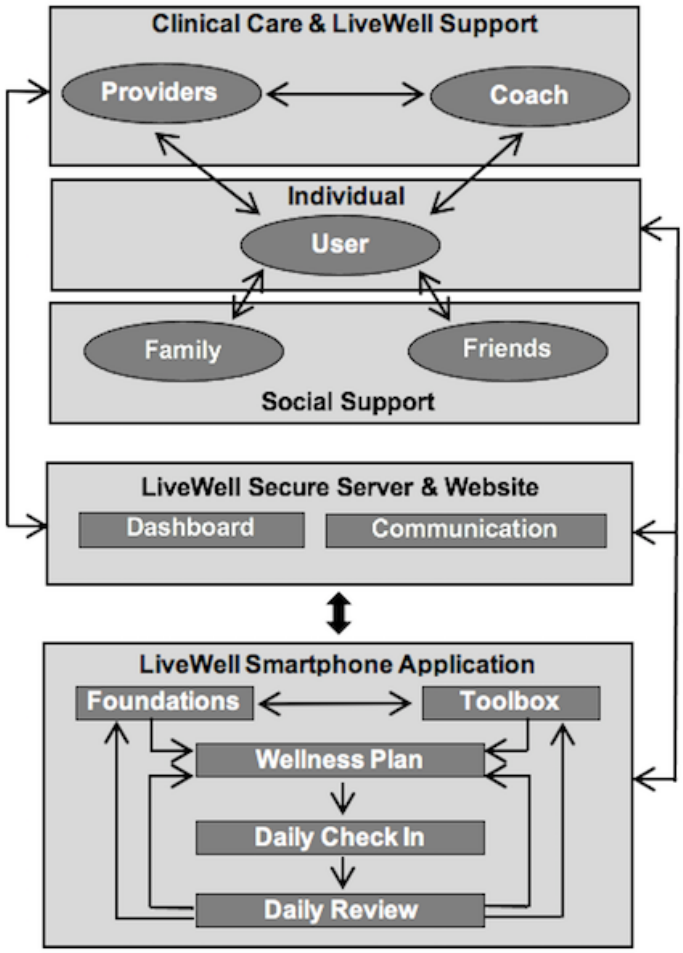
Intervention Overview

Similar to existing, empirically supported F2F psychotherapy interventions for bipolar disorder, the LiveWell intervention aims to decrease episode relapse, reduce symptom burden, and improve quality of life [3,4,10-15,58-60]. The intervention seeks to achieve these outcomes by assisting individuals with managing behaviors proposed to underlie the impact of existing therapies [7,11-13,15,18,19,31,61,62]. LiveWell thus emphasizes identifying early warning signs and symptoms of relapse, developing plans and monitoring for relapse, and then enacting and adjusting plans as needed. In addition, LiveWell engages users in a similar process to support taking medications as prescribed, obtaining adequate sleep duration, and maintaining regular routines. LiveWell also addresses strengthening social support, managing stressors, and engaging in healthy habits regarding diet, exercise, and substance use.

LiveWell consists of technological and human support components, including a smartphone app, a secure server, a website, and coaching support (Figure 1). The smartphone app consists of 5 components: Foundations, Toolbox, Wellness Plan, Daily Check-in, and Daily Review. It provides foundational information on bipolar disorder self-management (Foundations) and a toolbox with self-assessment surveys and skills practice (Toolbox). The Foundations and Toolbox components support developing a personalized Wellness Plan to reduce relapse risk and manage signs and symptoms. The core of the intervention is a Daily Check-in, in which users monitor medication adherence, sleep duration, routine, and wellness. Based on Daily Check-in data, the Daily Review provides tailored feedback that directs users to relevant psychoeducation content (eg, using lifestyle skills to reduce relapse risk or coping skills for managing early warning signs and symptoms).

LiveWell also utilizes human support, a coach to improve app use adherence, self-management, and communication with mental health providers [63]. The app provides data summaries and alerts via a secure server and website to help the coach provide support and facilitate communication with mental health care providers. Personnel without professional mental health training provide coaching support to reduce costs and increase access [64,65]. A clear division of labor between the technology and the coach ensures that coaches operate within the scope of nonclinical practice. The technology functions as the psychotherapeutic strategy expert that provides status summaries and alerts to the coach, who uses structured scripts and flowsheets to serve as a technology use concierge.

Figure 1. Intervention overview.



Field Trial

Monitoring is a major determinant of behavior change [41] and an essential strategy of empirically supported bipolar disorder psychotherapies [3,11,25,60]. Individuals with bipolar disorder are interested in using self-monitoring tools [30,31]. Thus, the initial development of LiveWell focused on creating a smartphone-based self-monitoring tool for tracking moods, thoughts, sleep, and routine. In particular, the ability to distinguish between different levels of wellness (eg, doing well or responding as expected to events versus experiencing early warning signs of an episode) may be essential to staying well [33].

Due to the potential importance of this ability, the field-trial Daily Check-in included 7-point scales to monitor mood and thoughts (Figure 2, Table 1). During the app training session, coaches helped users establish mood and thought anchors (ie, words describing their prior experiences at different wellness levels) to make these scales more personally relevant and useful. However, coaches noticed that users' mood and thought anchors often overlapped (eg, “planning for my future” as a mood anchor; Table 1), and some anchors appeared to reflect behaviors, physical symptoms, and changes in motivation (eg, “sluggish” as a thought anchor; Table 1). In addition, users reported that the 7-point range was too restrictive (eg, “If people are more depressed, it would probably be useful to have a wider scale to describe it”). As a result, the mood and thought rating scales were collapsed into a single 9-point wellness rating scale (Figure 2).

Users were also given the option to complete the Daily Check-in multiple times a day, but coaches noted that allowing multiple check-ins did not appear to elicit the reflection needed to identify different wellness levels. Instead, the Daily Check-in seemed to be capturing momentary reactions to daily hassles and uplifts [66]. To encourage users to engage in reflective monitoring rather than in-the-moment rating, the Daily Check-in was restricted to allow only one check-in per day, and the coaching scripts were adjusted to encourage reflection on wellness status using the personalized anchors.

Figure 2. Daily Check-in development. (A) Field trial; (B) Design sessions; (C) Pilot study, version 1; (D) Pilot study, version 2.

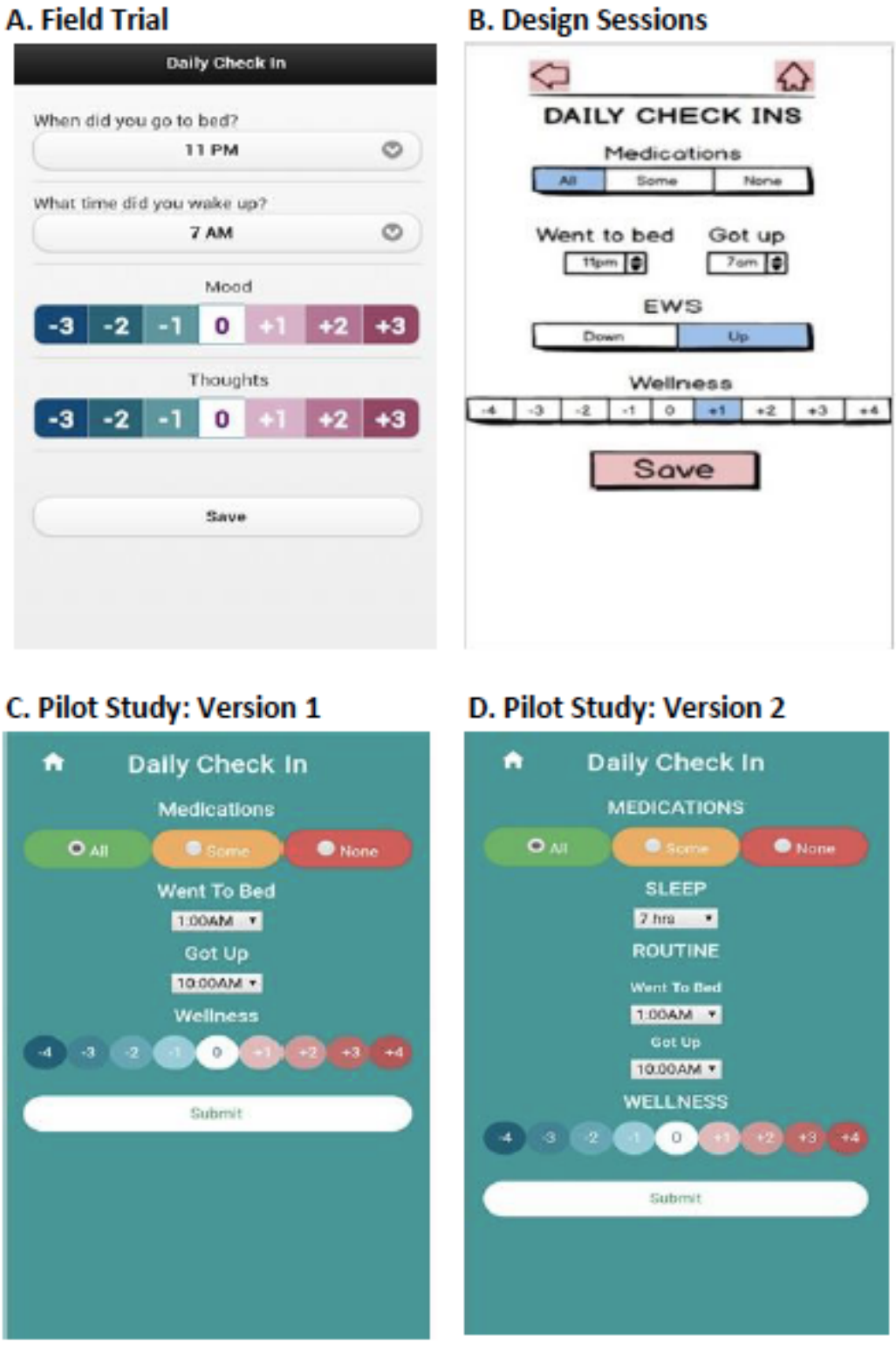


Table 1. Field trial: mood and thought rating scales (n=4).

Rating	Definition	Personalized anchor (user ID)	
		Mood anchor	Thought anchor
+3 Severe up	Most manic for you; extremely problematic	Ecstatic (2001) Uber social (2002) Sleep less than 12 hours (2008)	Racing (2001) Can work very fast (2002) Nonsensical (2003)
+2 Moderate up	Moderately problematic	Overly hopeful (2001) Full energizer (2008)	Feel up (2002) Have energy to study (2008)
+1 Slight up	Normal variation; explainable by recent or upcoming events	Anticipation (2001) Life is manageable (2002) Talkative (2003)	Realistic (2001) Feel competent (2002) Can study and enjoy it (2008)
0 Well	Doing well	Satisfied (2001) Cheerful (2003) Planning for my future (2008)	Reasonable (2001) Steady (2003) Learn new things (2008)
-1 Slight down	Normal variation; explainable by recent or upcoming events	Disappointed (2001) All or nothing kind of thinking (2002) Tired (2003)	Sluggish (2001) Normal thought rate (2003) Anxious (2003)
-2 Moderate down	Moderately problematic.	Cry easily (2002) Tired (2003).	Hopeless (2001) Don't feel like myself (2002)
-3 Severe down	Most depressed for you; extremely problematic	Despondent (2001) Emotional pain (2002) Overly lethargic (2003).	Despair (2001) Overly critical (2003) I can't study (2008)

Design Sessions and Usability Testing

Content and tools for the app and coaching support were developed based on information from empirically supported psychotherapies for bipolar disorder [3,12,13,15], health psychology behavior change theories [39,40,67-74], and chronic disease self-management models [75-83]. Design sessions and usability testing were then conducted to obtain user feedback on the overall app design, Daily Check-in, Daily Review, Foundations lessons, and the F2F coaching app training session.

During the design sessions, the Daily Check-in had separate early warning sign buttons to emphasize the importance of early warning signs (Figure 2). User feedback indicated these buttons were unnecessary and confusing with the transition to the 9-point rating scale. Thus, the Daily Check-in was simplified by removing these buttons, and early warning signs were clearly defined as +2 and -2 on the wellness rating scale (Table 2). Additionally, the Foundations content and coaching scripts were updated to emphasize the use of the wellness rating scale for early warning signs recognition.

During usability testing, user feedback suggested that exercise and diet were not adequately addressed. Therefore, Foundations and Toolbox content was added to cover the importance of exercise and diet [84-86]. User feedback also suggested that finding the right doctor and establishing a therapeutic alliance was not adequately discussed and that the rationale and approach to identifying a hospital in the case of a severe episode were unclear. As a result, the team lesson about working with a psychiatrist was expanded, the rationale for learning about a hospital for inpatient care was clarified, and information about using mental health directives was added.

In their posttask questionnaire responses, users indicated that the app's overall design and organization were straightforward, easy to use and understand, reasonable in terms of time commitment, and met their expectations (Table 3, Multimedia Appendix 5). Users also indicated that both the Foundations and Daily Review were interesting, relevant, and taught them something new.

Table 2. Design sessions: wellness rating scale (n=5).

Rating	Definition	Personalized wellness anchor (user ID)
+4 Severe up	Poor judgment, dangerous behaviors, not sleeping, hallucinations/delusions	Spend too much money (2002) Not sleeping (2003) Breaking stuff (2005)
+3 Moderate up	Many symptoms day to day, manic episode probably happening, difficult to maintain activities/routine	Physically energized (2001) Everything amplified (2002) Cussing out strangers (2005)
+2 Mild up	Some ongoing symptoms or early warning signs; manic episode may be coming, can still maintain activities/routine	Mood more volatile (2002) Binge drinking (2005) Walking for a long distance (2008)
+1 Slight up	Response to recent/upcoming good event, likely normal variation in wellness, understandable and manageable	More hopeful (2002) Chipper (2005) Exercise a little bit (2008)
0 Balanced	Neither up nor down, doing well	Engaged in life (2002) Like being around people (2005) Mood happy (2008)
-1 Slight down	Response to recent/upcoming bad event, likely normal variation in wellness, understandable and manageable	I'd rather sit at home (2001) Slight upset (2003) Sarcastic (2005)
-2 Mild down	Some ongoing symptoms or early warning signs, depressive episode may be coming, can still maintain activities/routine	Angry (2002) Selectively returning messages (2005) Loss of energy level interferes with daily tasks (2008)
-3 Moderate down	Many symptoms day to day, depressive episode probably happening, difficult to maintain activities/routine	Using sleep to avoid life (2001) Skipping meals (2005) Slow thought process (2008)
-4 Severe down	Serious ideas about suicide, immobilized, dangerous behaviors, disrupted sleep, hallucinations/delusions	Lack of motivation to do things (2003) Suicide planning (2005) Loss of interest in everything (2008)

Table 3. Usability testing: posttask questionnaire (n=5).

Section and usability type	Question	DSD[a]	ASA[b]
Overall			
Interface quality	This application is visually appealing.	0	80
Interface quality	It was easy to move from one page to another.	0	80
Ease of learning	The overall organization of the application is easy to understand.	0	100
Interface quality	Individual pages are well designed.	0	80
Ease of learning	Terminology used in this application is clear.	0	80
Satisfaction	The content of the application met my expectations.	0	60
Satisfaction	I would be likely to use this application in the future.	0	80
Ease of use	I was able to complete my tasks in a reasonable amount of time.	0	80
Ease of use	Overall, the application is easy to use.	0	100
Foundations			
Ease of learning	I found the lessons easy to understand.	0	80
Usefulness	I found the lessons interesting.	0	80
Usefulness	I found the lessons relevant to me.	0	60
Usefulness	I learned something new from the lessons.	0	60
Usefulness	I was motivated to make a change after reading the lessons.	0	60
Daily Review			
Ease of learning	I found the daily review easy to understand.	0	80
Usefulness	I found the daily review interesting.	0	80
Usefulness	I found the daily review relevant to me.	0	60
Usefulness	I learned something new from the daily review.	0	80
Usefulness	I was motivated to make a change after completing the daily review.	0	80

[a]DSD: disagree/strongly disagree.

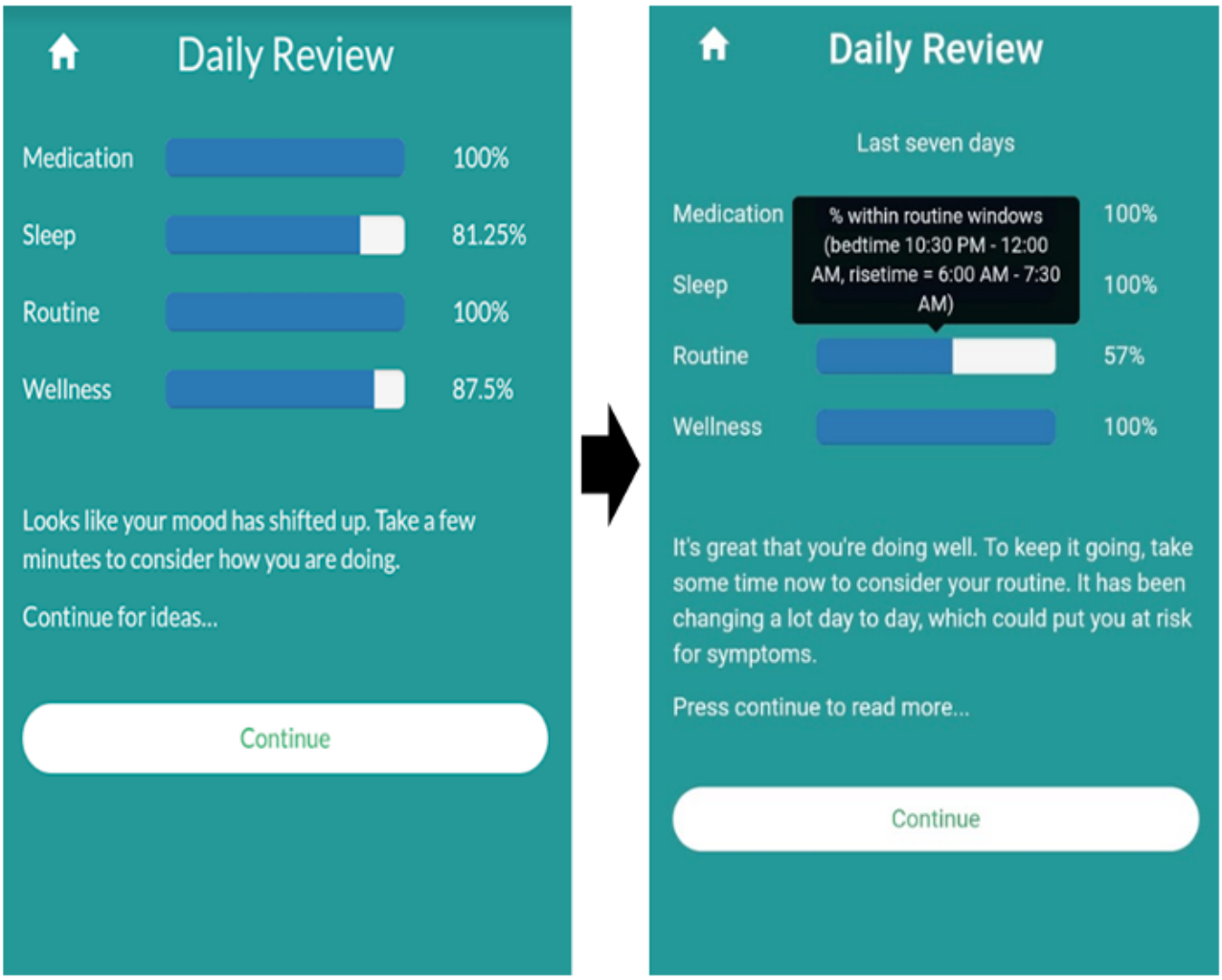
[b]ASA: agree/strongly agree.

Pilot Study

The pilot study tested the complete intervention, including the technology and human support. At the F2F session, coaches helped users identify personalized wellness anchors (Table 2) and set mutually agreed-upon target goals (medication adherence, sleep duration, routine bedtime and rise time, and wellness rating range) for daily monitoring (Figure 1). The coaches encouraged users to set goals known to facilitate staying well [11-13,87], such as taking their medications 100% of the time, sleeping the recommended amount each day (7-9 hours, with 6-10 hours reasonable for some), going to bed and starting their day within a 1.5-hour window, and keeping their wellness ratings within a balanced range (-1/+1; Table 2).

Users found differentiating between maintaining a regular routine and getting the right amount of sleep confusing. Therefore, a separate button for recording sleep duration was added to the Daily Check-in (Figure 1), and coaching scripts were updated to clarify the behavioral targets and goal setting process (Multimedia Appendix 6). Users also expressed difficulties understanding the Daily Review feedback regarding goal achievement. The first page of Daily Review feedback uses bar graphs to display the percent of days over the last 7 days that the user met their target goal (Figure 3). A title was added to the Daily Review bar graphs to indicate that the percentage bars correspond to goal achievement over the last 7 days. A hover bar feature was also added so that users could view their personalized goals for each behavioral target (Figure 3). Some users found the Daily Review feedback repetitive when target goals were consistently not met. To allow users to sample additional Daily Review content under these conditions, a Review Something Else component was added and made available after completing the suggested Daily Review feedback. Coaches found that users did not always follow up on Daily Review feedback to contact their providers. To address this, when feedback prompting communication with providers was given, a pop-up message with a link to one's psychiatrist's phone number was added to the Daily Review. Email alerts to coaches were also added, and coaching scripts were developed to help the coaches contact the user and psychiatrist under these conditions [63].

Figure 3. Daily Review development.



At week 4, users were asked to complete reading the Foundation lessons, and coaches assisted users in developing a lifestyle plan to reduce relapse risk (Table 4) and a coping plan for managing signs and symptoms (Table 5). Originally, generic suggestions for the plans were included. However, when coaches asked users to personalize their plans, some wanted to keep the generic plan. Generic plans were removed for the diet, exercise, substances (Attend), stress management (Tranquil), and build-support (Social) sections of the Reduce Risk plan to encourage users to think of relevant personal plans (Multimedia Appendix 7). Users' personalized goals for sleep duration and bedtime/rise time windows, and a generic recommendation to take all of their psychiatric medications daily, were retained. Coaching scripts were adapted to encourage the use of implementation intentions for planning [88-90], and users were given examples of implementation intentions in the generic risk-reduction plans (eg, “If you miss medications for more than 3-4 days in a row, then discuss with your psychiatrist”).

Table 4. Pilot study: lifestyle plan for reducing risk (n=11).

Target	Personalized plan (user ID)
Sleep	Keep routine. (2001)
	Sleep diary if not going well (2005)
	If having trouble sleeping, try sleep medication. (2066)
Medication	Talk to psychiatrist about current medication. (2003)
	Solve how to get around barriers or hurry/inattentiveness. (2005)
	Take every medication at the same time every day (2041)
Attend	Pay attention to sugar and processed foods. (2065)
	Do yoga/tai chi at least 3 times a week. (2066)
	Limit caffeine to 1 cup of coffee a day. (2086)
Routine	Don't skip meals. (2003)
	Exercise at a consistent time. (2008)
	If having trouble getting up, set an alarm. (2063)
Tranquil	Do a crossword puzzle. (2016)
	Listen to relaxation strategies. (2041)
	Don't over-schedule to avoid too much stimulation. (2066)
Social	Go to work with a positive attitude. (2001)
	Explore churches. (2066)
	Socialize with friends or family at least twice a week. (2086)

Table 5. Pilot study: coping plan for managing signs and symptoms (n=11).

Rating	Definition	Personalized anchor and plan (user ID)	
		Anchor	Plan
+4 Severe up	Poor judgment, dangerous behaviors, not sleeping, hallucinations/delusions	Super aggressive (2005) Invincible (2001) Risky activities (2041)	Increase supervision. (2063) Let supports know. (2086)
+3 Moderate up	Many symptoms day to day, manic episode probably happening, difficult to maintain activities/routine	Hardly sleeping (2001) Really angry (2005) Excessive energy (2008)	Stick to routine. (2016) Lean on family for support. (2008)
+2 Mild up	Some ongoing symptoms or early warning signs, manic episode may be coming, can still maintain activities/routine	Bizarre thoughts (2016) Short-tempered (2005) Sleep less (2001)	Be conscious, notice signs. (2003) Ask supports if noticing manic symptoms. (2005)
+1 Slight up	Response to recent/upcoming good event, likely normal variation in wellness, understandable and manageable	Friendly (2005) More social (2016) More active (2041)	Watch substance intake, especially alcohol, during the week. (2063) Moderate alcohol consumption. (2001)
0 Balanced	Neither up nor down, doing well	Balanced sleep (2016) Energy level is normal (2001) Productive (2001)	Recognize you are doing a great job! (2041) Eat healthy and exercise. (2063)
-1 Slight down	Response to recent/upcoming bad event, likely normal variation in wellness, understandable and manageable	Pessimistic reaction to events (2005) Can't motivate self (2016) Irritable (2003)	Let self feel sad, then move on because [I] can't change events. (2008) Do things that make you happy. (2003)
-2 Mild down	Some ongoing symptoms or early warning signs, depressive episode may be coming, can still maintain activities/routine	Crying (2016) Start losing appetite (2005) More restless (2008)	Try not to sleep too much or spend too much time in bed. (2003) Live day by day. (2008)
-3 Moderate down	Many symptoms day to day, depressive episode probably happening, difficult to maintain activities/routine	Trouble leaving house (2041) Oversleeping (2016) Grooming not shaving (2008)	Go to [my] mom's house. (2016) Call family supports. (2041)
-4 Severe down	Serious ideas about suicide. Immobilized. Dangerous behaviors. Disrupted sleep. Hallucinations/delusions.	Actively trying to harm self (2003) Odd thoughts (2008) Hate cycle (2016)	See doctor immediately. (2008) Contact those you trust. (2003)

Users reported confusion and difficulties with setting and achieving target goals and did not consistently follow up with Daily Review recommendations to contact providers; thus, the coach role was expanded from app use support [65] to more active support for self-management and clinical care communication. Clear coach email alerts and dashboard summaries, scripts, and flowsheets were developed to support these additional roles [63]. The coaching scripts were modified to emphasize creating target behavioral goals, assisting users with developing plans, reinforcing success, and making adjustments when goals were not met (Multimedia Appendix 6). This increased self-management support included developing flowsheets and tip sheets to help coaches guide the user to appropriate self-management content [63]. Questions from coaches about how to address worsening symptoms resulted in additional structured protocols to assess suicidality and functional impairment, including clear instructions on when to call for real-time clinical support. Details of the rationale and implementation of the coaching support are described elsewhere [63].

Pilot Study Exit Analysis

When users finished the 8-week pilot study, they were asked to complete a usability questionnaire (Multimedia Appendix 9) [91,92]. Responses are summarized in Table 6 for questions in which ≥7/11 users rated an item as “strongly agree” or “agree,” or ≥3/11 users rated an item as “strongly disagree” or “disagree.” All users found the Daily Check-in easy to use, and most indicated that it made them more aware of their early warning signs and sleep duration. Users expressed that the Foundations were easy to understand, about the right length, and relevant. They reported that they liked being able to personalize the Wellness Plan and found it relevant for their continued use of the app. Finally, users emphasized the importance of coaching support. They found the coach supportive and the calls convenient in terms of their schedule and appropriate length. Most users indicated that they were motivated by the coach to review intervention content and change their behaviors after coach calls.

Table 6. Pilot study: usability questionnaire (n=11).

Section and usability type	Question	DSD[a]	ASA[b]
Overall			
Ease of learning	The terminology used in this application is clear.	0	91
Ease of use	I was able to complete my tasks in a reasonable amount of time.	0	91
Satisfaction	The content of the application met my expectations.	9	73
Foundations			
Ease of learning	I found the lessons easy to understand.	0	100
Ease of use	The lessons were about the right length.	0	82
Usefulness	I found the lessons relevant to me.	0	73
Wellness Plan			
Usefulness	I found the Wellness Plan relevant to me.	0	91
Ease of learning	I found the Wellness Plan easy to understand.	0	82
Usefulness	I liked being able to personalize the Wellness Plan.	0	82
Usefulness	I learned something new from using the Wellness Plan.	0	73
Usefulness	Having and using my personal Wellness Plan was useful for me.	0	73
Daily Check-in			
Ease of learning	I found the Daily Check-in easy to use.	0	100
Usefulness	Using the Daily Check-in made me more aware of how much I was sleeping.	0	91
Usefulness	Using the Daily Check-in made me more aware of symptoms and early warning signs.	9	82
Usefulness	I found using the Daily Check-in helpful.	0	73
Daily Review			
Ease of learning	I found the Daily Review easy to understand.	0	100
Coach			
Usefulness	I found the coach supportive.	0	91
Usefulness	I found the coach calls useful.	0	91
Ease of use	The coach calls were an appropriate length.	0	91
Ease of use	I was able to schedule the coach calls at times that were convenient for me.	0	91
Usefulness	I got more out of the application by working with the coach.	0	82
Usefulness	I found the coach's role beneficial to my use of the application.	0	82
Usefulness	Having the coach calls motivated me to read the lessons.	0	73
Usefulness	I was motivated to make a change after phone calls with the coach.	0	73
Psychiatrist			
Usefulness	Using LiveWell helped me communicate with my psychiatrist about how I was doing.	27	27
Technical			
Usefulness	Once I completed my daily LiveWell activities, the reminders stopped appearing.	27	36
Satisfaction	I found the reminders irritating.	45	27
Usefulness	The reminders came on schedule as I programmed them to.	36	18
Usefulness	I relied on the reminders to complete my daily LiveWell activities.	27	18
Ease of use	The battery life of the phone was adequate.	55	9
Ease of use	The study team was helpful and responsive to my technical issues.	0	82

[a]DSD: disagree/strongly disagree.

[b]ASA: agree/strongly agree.

Thematic analysis of exit interviews identified 6 usability themes: ease of use, ease of learning, usefulness, barriers, suggestions, and technical limitations (Multimedia Appendix 10). Users indicated that they found the app easy to use.

> *It was pretty clear what was going on. . . . The Daily Check-in was very easy. It took very little time. It [Daily Review] very quickly gave me information on what it thought was going well. (2005)*

Some users reported that having self-management resources on their phones was especially convenient when they experienced symptoms. "It's sort of like having a therapist in your phone" (2041).

Users found the Foundation lessons and Wellness Plan relevant and useful. They stated that they liked being able to personalize their wellness anchors and plans and that this process made their app use more meaningful.

> *We personalized [the Wellness Plan], going through each mood variation level and noting what I've personally experienced. At first, it was really generic, so going through it and having it be like "if you are up a level, this is what you are going to be seeing" . . . It's my plan, not just a generic one. (2016)*

Users also stated that personalizing their information inspired self-awareness.

> *Personalizing the information was really helpful, within the Wellness Plan, within triggers . . . It made me so much more aware of myself. (2066)*

Moreover, users felt that the coaching support was motivating, useful, and helped them get more out of their app use.

> *I found the coaching really helpful. . . . [It] was helpful for motivation and for answering questions because the app was confusing at the beginning. (2061)*

Users offered many suggestions about changes or additions to data viewing, navigation, personalization, and monitoring, such as adding tracking for energy level, alcohol use, diet, and exercise. "There should be more visual things . . . like video clips" (2063).

> *The one missing [from the Daily Check-in] is energy level. It's very critical for bipolar. (2008)*

Users noted the app's navigation could be complicated, which sometimes led to not using parts of the app; however, coaching support helped resolve navigation difficulties. "Navigating was a little hard to get used to. . . . I'm not sure I ever looked at [the toolbox]" (2041).

> *It [Navigation] got easier. Like I didn't get at first how you could put things in my toolbox, and I mean [my coach] explained it all to me. (2086)*

Users still found that the Daily Review feedback could be repetitive if they were consistently experiencing problems with one specific behavioral target (eg, routine). In some cases, this repetitive feedback led to a discontinuation of use. "I got fed up after a couple weeks 'cause [the Daily Review] was the same thing" (2063).

Some users felt that technical issues impeded their use of the app. Specifically, users experienced problems with the reminders to check in, which contributed to irritation and reduced reliance on the reminders.

> *Sometimes I can't complete the [Daily Check-in] until later in the day, so pops up, I hit yes . . . and then it keeps popping up until it goes through all the reminder times that you missed. (2003)*

Other users expressed difficulties with study equipment, such as frustration with poor phone battery life. "I had some difficulty with the phone. [It] would die all the time" (2016).

However, most users felt that the study team was helpful and responsive in resolving technical issues.

> *[The] technical things I brought up, [the coach] was like, "That's been a problem we are working on it." It's good to know that it's the phone and not me. (2005)*

Anchors

A thematic analysis of users' mood, thoughts, and wellness rating scale anchors was completed to explore what types of signs and symptoms are relevant for monitoring wellness. The anchors for the mood, thought, and wellness rating scales were coded into subtypes and types (Multimedia Appendix 11). When the field trial anchors were formally coded, only 46% of the mood scale's anchors were coded as "mood," while the remainder were coded as "cognition," "physical," "behavior," or "motivation." For the thought scale, 69% of the anchors were coded as "cognition," with the remainder coded mostly as "mood" and "behavior." In contrast, when users identified personal anchors for the pilot study wellness scale (Table 7), the anchor types were coded as "behavior" (38%), followed by "cognition" (28%), "mood" (17%), "physical," (10%) and "motivation" (7%). In terms of subtypes, anchors for thought content (17%), sleep (12%), thought process (10%), negative mood (10%), energy (9%), and social interactions (8%) accounted for two-thirds of the subtypes.

Table 7. Pilot study: thematic analysis of personalized wellness rating scale anchors (n=11).

Type (%) and subtype (%)	Personalized anchor (user ID)
Behavior: visible activities or timing of such activities (38)	
Sleep: quality, duration, timing, need (12)	Stay up all night (2001) Oversleeping (2003) Nocturnal (2016)
Social: nonaggressive interactions with other people (8)	Enjoy seeing people (2005) More social (2016) Avoiding family & friends (2041)
Risky: increasing risk of injury or harm (4)	Buying things I don't need (2003) Running traffic lights (2008) Driving too fast (2041)
Self-care: eating, drinking, grooming, hygiene, medications (4)	Well-fed (2001) Grooming, not shaving (2008) Forgetting meds (2016)
Leisure: for relaxation or enjoyment, including over engagement (3)	Making art (2003) Watching movies (2008) Exercising, walking (2041)
Speech: rate, rhythm, or volume of speech (3)	Less talkative (2016)
Work: employment, school, home care, volunteering (2)	Productive (2001) Study for 3-5 hours (2008) Going to work and school (2016)
Aggression: physical or psychological harm to person, object, or self (1)	Actively trying to harm self (2003) Super aggressive (2005) Self-harm (2016)
Substance: ingestion of psychoactive substances (1)	Drinking to dangerous excess (2001)
Cognition: acquiring knowledge and understanding through thought, experience, and senses (28)	
Content: what one is thinking about (17)	Life is not worth living (2001) Have lots of new ideas (2003) Odd thoughts (2008)
Process: logic, organization, coherence, and speed of thinking (10)	Faster thinking (2003) Bad judgment (2005) Problems with any decision (2041)
Perception: sensory processing, disturbances of sensory processing (1)	Hallucinations (2003)
Mood: emotional or affective state (17)	
Negative: unpleasant, disagreeable, lack of pleasure (10)	Irritable (2003) Persistent crabbiness (2005) Sad mood, "feeling down" (2041)
Positive: good, affirmative, or constructive (6)	Things are so exciting (2001) Laugh (2005) Easier to smile (2041)
Physical: relating to the body (10)	
Energy: strength and vitality (9)	Feel fatigued (2001) More restless (2008) Restless (2016)
Appetite: desire for food (1)	Not hungry (2001) Force self to eat (2005)
Motivation: reasons or drive to engage in behavior (7)	

Type (%) and subtype (%)	Personalized anchor (user ID)
N/A[a]	Willing to try anything (2001)
	Loss of interest (2008)
	Easier to maintain a routine (2041)

[a]N/A: not applicable.

Discussion

This paper describes the user-centered process that guided the development of the technological and human support components of LiveWell. Consistent with the importance of monitoring [3,11,25,60], the core of the LiveWell intervention consists of a daily self-monitoring tool, the Daily Check-in, which underwent multiple revisions during development. Changes to the technology and the coaching support were also made to help users set clear target goals, track the achievement of those goals, make adjustments, and communicate with care providers.

The initial Daily Check-in included both mood and thought rating scales. Coaches helped users develop personalized anchors for these scales to help users identify early warning signs and symptoms of an episode based on their past experiences. Examining these anchors revealed that the users' anchors often did not fit within the bounds of the requested mood and thought categories. This suggested that users were being asked to personalize scales that did not capture their most relevant wellness variation experiences. Therefore, the mood and thought scales were collapsed into a single wellness rating scale. When users were provided with more freedom to identify personal anchors using the wellness scale, the anchors most commonly reflected behavioral—followed by cognitive, mood, physical, and motivational—signs and symptoms. More specifically, two-thirds of these anchors referenced thought content and process, sleep, negative mood, energy, and social interactions. These findings are consistent with the early warning sign literature indicating that individuals typically endorse changes in cognition (concentration, self-esteem, difficulty with decisions), behavior (more talkative and aggressive, changes in sleep duration), and energy before episodes [93,94].

Additional changes to the Daily Check-in were made to help users distinguish between getting adequate sleep and keeping a regular routine. Users reported that the Daily Check-in made them more aware of early warning signs and symptoms and how much they were sleeping. Early warning sign management and sleep duration are considered to be important targets that underlie the improved outcomes produced by empirically supported therapies for bipolar disorder [3,11,25,60]. It thus appears that user feedback led to changes in the Daily Check-in that may assist users in staying well.

Changes to the technology were also made to clarify goal setting, goal achievement tracking, and making adjustments (Daily Check-in and Daily Review), as users experienced difficulties with these strategies. In addition, changes were made to the Foundations and Toolbox to clarify the process of setting goals, making plans, monitoring, evaluating goal achievement, and making adjustments if needed. Due to the inconsistency with which users would act upon recommendations to contact their provider, pop-up notifications with a link to their psychiatrist's phone number were added.

However, it was unclear that these technology changes were sufficient, so coaching roles were expanded to include support for self-management strategy use and clinical care communication. Research suggests that human support for app use reduces attrition and improves adherence; however, increased engagement does not always translate into improved outcomes [29,38,66,95-97]. Improved outcomes may arise from the inclusion of clinical support to ensure that users identify the content and tools relevant to their needs, use them correctly, and translate this use into their daily lives [38]. This suggests that expanding the coaching roles for LiveWell may improve outcomes. In creating the self-management and clinical care communication support roles for LiveWell, feedback from the coaches played an important role in developing these roles.

Users reported that working with their coaches to tailor their wellness rating scale anchors and plans made the intervention more relevant to their personal experiences. This feedback is consistent with prior studies indicating that personalizing application components to address user needs can increase user engagement and positively impact intervention outcomes [43,98,99]. Thus, the provision of generic plans was minimized to encourage the personalization of the wellness plans. Users noted that developing personally relevant plans motivated them to enact these plans. Taken together, these findings indicate that self-management interventions that utilize open-ended, personalized wellness scales and plans may help individuals develop insight into their health condition and more readily embrace and act on intervention content.

Finally, striking a balance between making the app easy to navigate and fulfilling participant requests for additional features was challenging. As user engagement typically decreases with challenging-to-use applications [100], an effort was made to prioritize user requests related to intervention functionality. Technical issues, such as problems with reminders and battery life, have been addressed with improvements in technology over time, such as the integration of app reminders into the Android operating system and smartphone battery-life improvements.

To support the ongoing improvement and dissemination of technology-based mental health interventions [38-42], we have provided a detailed description of the user-centered development process for LiveWell. This process suggests that individuals with bipolar disorder value target monitoring, personalization of goals and plans, and human support aids as self-management tools. In developing LiveWell's technology and human support, feedback from both users and coaches played an important role, emphasizing the significance of engaging all stakeholders in

intervention development. This attention to all-stakeholder input is broadly applicable to developing the technology and the human support for digital mental health interventions.

Conflicts of Interest

DCM has accepted honoraria and consulting fees from Apple Inc, Otsuka Pharmaceuticals, Pear Therapeutics, and the One Mind Foundation; royalties from Oxford Press; and an ownership interest in Adaptive Health Inc. EHG has accepted honoraria from Otsuka Pharmaceuticals. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Field trial: scripts handouts.
[PDF File (Adobe PDF File), 277 KB - mental_v8i4e20424_app1.pdf]

Multimedia Appendix 2
Design sessions: coaching script handouts.
[PDF File (Adobe PDF File), 287 KB - mental_v8i4e20424_app2.pdf]

Multimedia Appendix 3
Design sessions: smartphone app mock-up.
[PDF File (Adobe PDF File), 551 KB - mental_v8i4e20424_app3.pdf]

Multimedia Appendix 4
Usability testing: script scenarios.
[PDF File (Adobe PDF File), 210 KB - mental_v8i4e20424_app4.pdf]

Multimedia Appendix 5
Usability testing: posttask questionnaire responses.
[XLSX File (Microsoft Excel File), 14 KB - mental_v8i4e20424_app5.xlsx]

Multimedia Appendix 6
Pilot study: coaching scripts handouts.
[PDF File (Adobe PDF File), 516 KB - mental_v8i4e20424_app6.pdf]

Multimedia Appendix 7
Pilot study: user personalized plans.
[XLSX File (Microsoft Excel File), 31 KB - mental_v8i4e20424_app7.xlsx]

Multimedia Appendix 8
Pilot study: exit interview script.
[PDF File (Adobe PDF File), 116 KB - mental_v8i4e20424_app8.pdf]

Multimedia Appendix 9
Pilot study: exit questionnaire responses.
[XLSX File (Microsoft Excel File), 22 KB - mental_v8i4e20424_app9.xlsx]

Multimedia Appendix 10
Pilot study: exit interview thematic analysis.
[XLSX File (Microsoft Excel File), 120 KB - mental_v8i4e20424_app10.xlsx]

Multimedia Appendix 11
User personalized anchors thematic analysis.
[XLSX File (Microsoft Excel File), 1852 KB - mental_v8i4e20424_app11.xlsx]

References

1. Judd LL, Akiskal HS, Schettler PJ, Endicott J, Maser J, Solomon DA, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. Arch Gen Psychiatry 2002 Jun;59(6):530-537. [doi: 10.1001/archpsyc.59.6.530] [Medline: 12044195]
2. Judd LL, Schettler PJ, Akiskal HS, Coryell W, Leon AC, Maser JD, et al. Residual symptom recovery from major affective episodes in bipolar disorders and rapid episode relapse/recurrence. Arch Gen Psychiatry 2008 Apr;65(4):386-394. [doi: 10.1001/archpsyc.65.4.386] [Medline: 18391127]
3. Miklowitz DJ. Adjunctive psychotherapy for bipolar disorder: state of the evidence. Am J Psychiatry 2008 Nov;165(11):1408-1419 [FREE Full text] [doi: 10.1176/appi.ajp.2008.08040488] [Medline: 18794208]
4. Crowe M, Whitehead L, Wilson L, Carlyle D, O'Brien A, Inder M, et al. Disorder-specific psychosocial interventions for bipolar disorder--a systematic review of the evidence for mental health nursing practice. Int J Nurs Stud 2010 Jul;47(7):896-908. [doi: 10.1016/j.ijnurstu.2010.02.012] [Medline: 20233615]
5. World Health Organization. Burden of Mental and Behavioural Disorders. In: The World Health Report 2001 - Mental health: new understanding, new hope. Geneva, Switzerland: World Health Organization; 2001.
6. Keck PE, McElroy SL, Strakowski SM, West SA, Sax KW, Hawkins JM, et al. 12-month outcome of patients with bipolar disorder following hospitalization for a manic or mixed episode. Am J Psychiatry 1998 May;155(5):646-652. [doi: 10.1176/ajp.155.5.646] [Medline: 9585716]
7. Geddes JR, Miklowitz DJ. Treatment of bipolar disorder. Lancet 2013 May 11;381(9878):1672-1682 [FREE Full text] [doi: 10.1016/S0140-6736(13)60857-0] [Medline: 23663953]
8. Miura T, Noma H, Furukawa TA, Mitsuyasu H, Tanaka S, Stockton S, et al. Comparative efficacy and tolerability of pharmacological treatments in the maintenance treatment of bipolar disorder: a systematic review and network meta-analysis. The Lancet Psychiatry 2014 Oct;1(5):351-359. [doi: 10.1016/s2215-0366(14)70314-1]
9. Goodwin GM, Consensus Group of the British Association for Psychopharmacology. Evidence-based guidelines for treating bipolar disorder: revised second edition--recommendations from the British Association for Psychopharmacology. J Psychopharmacol 2009 Jun;23(4):346-388. [doi: 10.1177/0269881109102919] [Medline: 19329543]
10. Gregory VL. Cognitive-behavioral therapy for depression in bipolar disorder: a meta-analysis. J Evid Based Soc Work 2010 Jul;7(4):269-279. [doi: 10.1080/15433710903176088] [Medline: 20799127]
11. Miklowitz DJ, Goodwin GM, Bauer MS, Geddes JR. Common and specific elements of psychosocial treatments for bipolar disorder: a survey of clinicians participating in randomized trials. J Psychiatr Pract 2008 Mar;14(2):77-85 [FREE Full text] [doi: 10.1097/01.pra.0000314314.94791.c9] [Medline: 18360193]
12. Miklowitz DJ, Scott J. Psychosocial treatments for bipolar disorder: cost-effectiveness, mediating mechanisms, and future directions. Bipolar Disord 2009 Jun;11 Suppl 2:110-122. [doi: 10.1111/j.1399-5618.2009.00715.x] [Medline: 19538690]
13. Scott J, Colom F, Vieta E. A meta-analysis of relapse rates with adjunctive psychological therapies compared to usual psychiatric treatment for bipolar disorders. Int J Neuropsychopharmacol 2007 Feb;10(1):123-129 [FREE Full text] [doi: 10.1017/S1461145706006900] [Medline: 16787554]
14. Lam DH, Burbeck R, Wright K, Pilling S. Psychological therapies in bipolar disorder: the effect of illness history on relapse prevention - a systematic review. Bipolar Disord 2009 Aug;11(5):474-482. [doi: 10.1111/j.1399-5618.2009.00724.x] [Medline: 19624386]
15. Schöttle D, Huber CG, Bock T, Meyer TD. Psychotherapy for bipolar disorder: a review of the most recent studies. Curr Opin Psychiatry 2011 Nov;24(6):549-555. [doi: 10.1097/YCO.0b013e32834b7c5f] [Medline: 21918448]
16. Olfson M, Marcus SC, Druss B, Pincus HA. National trends in the use of outpatient psychotherapy. Am J Psychiatry 2002 Nov;159(11):1914-1920. [doi: 10.1176/appi.ajp.159.11.1914] [Medline: 12411228]
17. Brown JD, Barrett A, Hourihan K, Caffery E, Ireys HT. State Variation in the Delivery of Comprehensive Services for Medicaid Beneficiaries with Schizophrenia and Bipolar Disorder. Community Ment Health J 2015 Jul;51(5):523-534. [doi: 10.1007/s10597-015-9857-5] [Medline: 25786723]
18. Busch AB, Ling D, Frank RG, Greenfield SF. Changes in the quality of care for bipolar I disorder during the 1990s. Psychiatr Serv 2007 Jan;58(1):27-33 [FREE Full text] [doi: 10.1176/ps.2007.58.1.27] [Medline: 17215409]
19. Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):629-640. [doi: 10.1001/archpsyc.62.6.629] [Medline: 15939840]
20. Denizard-Thompson NM, Feiereisel KB, Stevens SF, Miller DP, Wofford JL. The digital divide at an urban community health center: implications for quality improvement and health care access. J Community Health 2011 Jun;36(3):456-460. [doi: 10.1007/s10900-010-9327-5] [Medline: 21086028]
21. Proudfoot J, Parker G, Hadzi PD, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010 Dec;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]
22. Ben-Zeev D, Davis KE, Kaiser S, Krzsos I, Drake RE. Mobile technologies among people with serious mental illness: opportunities for future services. Adm Policy Ment Health 2013 Jul;40(4):340-343 [FREE Full text] [doi: 10.1007/s10488-012-0424-x] [Medline: 22648635]

23. Torous J, Wisniewski H, Liu G, Keshavan M. Mental Health Mobile Phone App Usage, Concerns, and Benefits Among Psychiatric Outpatients: Comparative Survey Study. JMIR Ment Health 2018 Nov 16;5(4):e11715 [FREE Full text] [doi: 10.2196/11715] [Medline: 30446484]
24. Hollis C, Morriss R, Martin J, Amani S, Cotton R, Denis M, et al. Technological innovations in mental healthcare: harnessing the digital revolution. Br J Psychiatry 2015 Apr 02;206(4):263-265. [doi: 10.1192/bjp.bp.113.142612] [Medline: 25833865]
25. Lam D, Wong G. Prodromes, coping strategies and psychological interventions in bipolar disorders. Clin Psychol Rev 2005 Dec;25(8):1028-1042. [doi: 10.1016/j.cpr.2005.06.005] [Medline: 16125292]
26. Lam D, Bright J, Jones S, Hayward P, Schuck N, Chisholm D. Cognitive therapy for bipolar illness : a pilot study of relapse prevention. Cognitive Therapy and Research 2000 Oct;24(5):503-520. [doi: 10.1023/A:1005557911051]
27. Lam DH, Hayward P, Watkins ER, Wright K, Sham P. Relapse prevention in patients with bipolar disorder: cognitive therapy outcome after 2 years. Am J Psychiatry 2005 Feb;162(2):324-329. [doi: 10.1176/appi.ajp.162.2.324] [Medline: 15677598]
28. Meyer TD, Hautzinger M. Cognitive behaviour therapy and supportive therapy for bipolar disorders: relapse rates for treatment period and 2-year follow-up. Psychol. Med 2011 Nov 21;42(7):1429-1439. [doi: 10.1017/s0033291711002522]
29. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Aug;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
30. Michalak EE, Suto MJ, Barnes SJ, Hou S, Lapsley S, Scott MW, Crest Bd. Effective self-management strategies for bipolar disorder: A community-engaged Delphi Consensus Consultation study. J Affect Disord 2016 Dec;206:77-86 [FREE Full text] [doi: 10.1016/j.jad.2016.06.057] [Medline: 27466745]
31. Murray G, Suto M, Hole R, Hale S, Amari E, Michalak EE. Self-management strategies used by 'high functioning' individuals with bipolar disorder: from research to clinical practice. Clin Psychol Psychother 2011 Apr;18(2):95-109. [doi: 10.1002/cpp.710] [Medline: 20572206]
32. Janney CA, Bauer MS, Kilbourne AM. Self-management and bipolar disorder--a clinician's guide to the literature 2011-2014. Curr Psychiatry Rep 2014 Sep;16(9):485. [doi: 10.1007/s11920-014-0485-5] [Medline: 25123130]
33. Russell SJ, Browne JL. Staying well with bipolar disorder. Aust N Z J Psychiatry 2005 Mar 26;39(3):187-193. [doi: 10.1080/j.1440-1614.2005.01542.x] [Medline: 15701069]
34. Nicholas J, Boydell K, Christensen H. Beyond symptom monitoring: Consumer needs for bipolar disorder self-management using smartphones. Eur Psychiatry 2017 Jul;44:210-216. [doi: 10.1016/j.eurpsy.2017.05.023] [Medline: 28692910]
35. Daus H, Kislicyn N, Heuer S, Backenstrass M. Disease management apps and technical assistance systems for bipolar disorder: Investigating the patients´ point of view. J Affect Disord 2018 Mar 15;229:351-357. [doi: 10.1016/j.jad.2017.12.059] [Medline: 29331693]
36. Murnane EL, Cosley D, Chang P, Guha S, Frank E, Gay G, et al. Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: implications for the design of technologies to manage mental health. J Am Med Inform Assoc 2016 May;23(3):477-484. [doi: 10.1093/jamia/ocv165] [Medline: 26911822]
37. Todd NJ, Jones SH, Lobban FA. What do service users with bipolar disorder want from a web-based self-management intervention? A qualitative focus group study. Clin Psychol Psychother 2013;20(6):531-543. [doi: 10.1002/cpp.1804] [Medline: 22715161]
38. Schueller SM, Tomasino KN, Mohr DC. Integrating Human Support Into Behavioral Intervention Technologies: The Efficiency Model of Support. Clin Psychol Sci Pract 2016 Nov 17;24(1):27-45. [doi: 10.1111/cpsp.12173]
39. Brendryen H, Kraft P, Schaalma H. Looking Inside the Black Box: Using Intervention Mapping to Describe the Development of the Automated Smoking Cessation Intervention ‘Happy Ending’. J Smok Cessat 2012 Feb 21;5(1):29-56. [doi: 10.1375/jsc.5.1.29]
40. Brendryen H, Johansen A, Nesvåg S, Kok G, Duckert F. Constructing a Theory- and Evidence-Based Treatment Rationale for Complex eHealth Interventions: Development of an Online Alcohol Intervention Using an Intervention Mapping Approach. JMIR Res Protoc 2013 Jan;2(1):e6 [FREE Full text] [doi: 10.2196/resprot.2371] [Medline: 23612478]
41. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychology 2008;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379]
42. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. American Journal of Preventive Medicine 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015]
43. Van Velsen L, Wentzel J, Van Gemert-Pijnen JE. Designing eHealth that Matters via a Multidisciplinary Requirements Development Approach. JMIR Res Protoc 2013 Jun 24;2(1):e21 [FREE Full text] [doi: 10.2196/resprot.2547] [Medline: 23796508]
44. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
45. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007 May 27;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]

46. 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]
47. Sachs GS, Guille C, McMurrich SL. A clinical monitoring form for mood disorders. Bipolar Disord 2002 Oct;4(5):323-327. [doi: 10.1034/j.1399-5618.2002.01195.x] [Medline: 12479665]
48. Sachs GS, Thase ME, Otto MW, Bauer M, Miklowitz D, Wisniewski SR, et al. Rationale, design, and methods of the systematic treatment enhancement program for bipolar disorder (STEP-BD). Biol Psychiatry 2003 Jun 01;53(11):1028-1042. [doi: 10.1016/s0006-3223(03)00165-3] [Medline: 12788248]
49. Joe J, Chaudhuri S, Le T, Thompson H, Demiris G. The use of think-aloud and instant data analysis in evaluation research: Exemplar and lessons learned. J Biomed Inform 2015 Aug;56:284-291 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.001] [Medline: 26071683]
50. Kjeldskov J, Skov M, Stage J. Instant data analysis: conducting usability evaluations in a day. In: NordiCHI '04: Proceedings of the third Nordic conference on Human-computer interaction. 2004 Presented at: NordiCHI04: NordiCHI 2004; October, 2004; Tampere Finland p. 233-240 URL: https://dl.acm.org/doi/abs/10.1145/1028014.1028050 [doi: 10.1145/1028014.1028050]
51. de Best MA, Deurzen VK, Geels M, van Oostrom F, Pries J, Scheele B. Quick and clean, fast and efficient usability testing for software redesign. 2009 Presented at: Proc Chi Ned; 2009; Delft, Netherlands p. 149-152 URL: https://www.frenspries.nl/quick%20and%20clean.pdf
52. Potter M, Gordon S, Hamer P. The nominal group technique: a useful consensus methodology in physiotherapy research. New Zealand Journal of Physiotherapy 2004;32:125-130.
53. Faulkner L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments, & Computers 2003 Aug;35(3):379-383. [doi: 10.3758/BF03195514]
54. Guest G. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. Field Methods 2006 Feb 01;18(1):59-82. [doi: 10.1177/1525822X05279903]
55. 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]
56. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. Journal of Usability Studies 2009;5(1):34-45.
57. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
58. Bond K, Anderson IM. Psychoeducation for relapse prevention in bipolar disorder: a systematic review of efficacy in randomized controlled trials. Bipolar Disord 2015 Jun;17(4):349-362. [doi: 10.1111/bdi.12287] [Medline: 25594775]
59. Gregory VL. Cognitive-Behavioral Therapy for Mania: A Meta-Analysis of Randomized Controlled Trials. Social Work in Mental Health 2010 Oct 07;8(6):483-494. [doi: 10.1080/15332981003744388]
60. Miklowitz DJ, Efthimiou O, Furukawa TA, Scott J, McLaren R, Geddes JR, et al. Adjunctive Psychotherapy for Bipolar Disorder: A Systematic Review and Component Network Meta-analysis. JAMA Psychiatry 2021 Feb 01;78(2):141-150. [doi: 10.1001/jamapsychiatry.2020.2993] [Medline: 33052390]
61. Reinares M, Sánchez-Moreno J, Fountoulakis KN. Psychosocial interventions in bipolar disorder: what, for whom, and when. J Affect Disord 2014 Mar;156:46-55. [doi: 10.1016/j.jad.2013.12.017] [Medline: 24439829]
62. Beynon S, Soares-Weiser K, Woolacott N, Duffy S, Geddes JR. Psychosocial interventions for the prevention of relapse in bipolar disorder: systematic review of controlled trials. Br J Psychiatry 2008 Jan;192(1):5-11. [doi: 10.1192/bjp.bp.107.037887] [Medline: 18174500]
63. Dopke C, McBride A, Babington P, Jonathan G, Michaels T, Ryan C. Development of Coaching Support for LiveWell: A Smartphone-Based Self-Management Intervention for Bipolar Disorder. JMIR Formative Research. Preprint posted online January 17, 2021. [doi: 10.2196/25810]
64. O'Mahen HA, Richards DA, Woodford J, Wilkinson E, McGinley J, Taylor RS, et al. Netmums: a phase II randomized controlled trial of a guided Internet behavioural activation treatment for postpartum depression. Psychol. Med 2013 Oct 09;44(8):1675-1689. [doi: 10.1017/s0033291713002092]
65. 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;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
66. 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] [Medline: 19403466]
67. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
68. Heaney CA. Intervention mapping and the new health promotion. Health Educ Behav 1998 Oct;25(5):564-8; discussion 569. [doi: 10.1177/109019819802500503] [Medline: 9768377]
69. Kok G, Mesters I. Getting inside the black box of health promotion programmes using intervention Mapping. Chronic Illn 2011 Sep;7(3):176-180. [doi: 10.1177/1742395311403013] [Medline: 21900338]

70. Kok G, Schaalma H, Ruiter RAC, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]
71. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
72. Cane J, Richardson M, Johnston M, Ladha R, Michie S. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. Br J Health Psychol 2015 Feb;20(1):130-150. [doi: 10.1111/bjhp.12102] [Medline: 24815766]
73. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: Mapping theoretically derived behavioural determinants to behaviour change techniques. Applied psychology 2008 Oct;57(4):660-680. [doi: 10.1111/j.1464-0597.2008.00341.x]
74. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
75. Von Korff M, Gruman J, Schaefer J, Curry SJ, Wagner EH. Collaborative management of chronic illness. Ann Intern Med 1997 Dec 15;127(12):1097-1102. [doi: 10.7326/0003-4819-127-12-199712150-00008] [Medline: 9412313]
76. Clark NM. Management of chronic disease by patients. Annu Rev Public Health 2003;24:289-313. [doi: 10.1146/annurev.publhealth.24.100901.141021] [Medline: 12415147]
77. Clark NM, Gong M, Kaciroti N. A model of self-regulation for control of chronic disease. Health Educ Behav 2001 Dec 30;28(6):769-782. [doi: 10.1177/109019810102800608] [Medline: 11720277]
78. Miller WR, Lasiter S, Bartlett Ellis R, Buelow JM. Chronic disease self-management: a hybrid concept analysis. Nurs Outlook 2015 Mar;63(2):154-161 [FREE Full text] [doi: 10.1016/j.outlook.2014.07.005] [Medline: 25241136]
79. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74(4):511-544. [Medline: 8941260]
80. Bartels SJ, Pratt SI, Mueser KT, Naslund JA, Wolfe RS, Santos M, et al. Integrated IMR for psychiatric and general medical illness for adults aged 50 or older with serious mental illness. Psychiatr Serv 2014 Mar 01;65(3):330-337 [FREE Full text] [doi: 10.1176/appi.ps.201300023] [Medline: 24292559]
81. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999 Jan;37(1):5-14. [doi: 10.1097/00005650-199901000-00003] [Medline: 10413387]
82. Salyers MP, Godfrey JL, McGuire AB, Gearhart T, Rollins AL, Boyle C. Implementing the illness management and recovery program for consumers with severe mental illness. Psychiatr Serv 2009 Apr;60(4):483-490. [doi: 10.1176/appi.ps.60.4.483] [Medline: 19339323]
83. Salyers MP, Godfrey JL, Mueser KT, Labriola S. Measuring illness management outcomes: a psychometric study of clinician and consumer rating scales for illness self management and recovery. Community Ment Health J 2007 Oct;43(5):459-480. [doi: 10.1007/s10597-007-9087-6] [Medline: 17514504]
84. Bauer IE, Gálvez JF, Hamilton JE, Balanzá-Martínez V, Zunta-Soares GB, Soares JC, et al. Lifestyle interventions targeting dietary habits and exercise in bipolar disorder: A systematic review. J Psychiatr Res 2016 Mar;74:1-7 [FREE Full text] [doi: 10.1016/j.jpsychires.2015.12.006] [Medline: 26724541]
85. Lopresti AL, Jacka FN. Diet and Bipolar Disorder: A Review of Its Relationship and Potential Therapeutic Mechanisms of Action. J Altern Complement Med 2015 Dec;21(12):733-739. [doi: 10.1089/acm.2015.0125] [Medline: 26348597]
86. Suto M, Murray G, Hale S, Amari E, Michalak EE. What works for people with bipolar disorder? Tips from the experts. J Affect Disord 2010 Jul;124(1-2):76-84. [doi: 10.1016/j.jad.2009.11.004] [Medline: 19969370]
87. Adams J, Scott J. Predicting medication adherence in severe mental disorders. Acta Psychiatr Scand 2000 Feb;101(2):119-124. [doi: 10.1034/j.1600-0447.2000.90061.x] [Medline: 10706011]
88. Gollwitzer PM. Implementation intentions: Strong effects of simple plans. American Psychologist 1999;54(7):493-503. [doi: 10.1037//0003-066x.54.7.493]
89. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement: A meta-analysis of effects and processes. Advances in Experimental Social Psychology 2006;38:69-119. [doi: 10.1016/S0065-2601(06)38002-1]
90. Sheeran P, Klein WMP, Rothman AJ. Health Behavior Change: Moving from Observation to Intervention. Annu Rev Psychol 2017 Jan 03;68:573-600. [doi: 10.1146/annurev-psych-010416-044007] [Medline: 27618942]
91. Lund A. Measuring Usability with the USE Questionnaire. Usability interface 2001;8(2).
92. Tullis T, Stetson JN. A Comparison of Questionnaries for Assessing Website Usability. 2004 Presented at: Usability professional association conference; June 7-11, 2004; Minneapolis, Minnesota. [doi: 10.1007/978-1-84628-795-4_3]
93. Altman ES, Rea MM, Mintz J, Miklowitz DJ, Goldstein MJ, Hwang S. Prodromal symptoms and signs of bipolar relapse: A report based on prospectively collected data. Psychiatry Research 1992 Jan;41(1):1-8. [doi: 10.1016/0165-1781(92)90012-r]
94. Keitner GI, Solomon DA, Ryan CE, Miller IW, Mallinger A, Kupfer DJ, et al. Prodromal and residual symptoms in bipolar I disorder. Comprehensive Psychiatry 1996 Sep;37(5):362-367. [doi: 10.1016/s0010-440x(96)90018-8]

95. Kelders SM, Bohlmeijer ET, Pots WT, van Gemert-Pijnen JE. Comparing human and automated support for depression: Fractional factorial randomized controlled trial. Behaviour Research and Therapy 2015 Sep;72:72-80. [doi: 10.1016/j.brat.2015.06.014]
96. Kleiboer A, Donker T, Seekles W, van Straten A, Riper H, Cuijpers P. A randomized controlled trial on the role of support in Internet-based problem solving therapy for depression and anxiety. Behaviour Research and Therapy 2015 Sep;72:63-71. [doi: 10.1016/j.brat.2015.06.013]
97. Mira A, Bretón-López J, García-Palacios A, Quero S, Baños RM, Botella C. An Internet-based program for depressive symptoms using human and automated support: a randomized controlled trial. NDT 2017 Mar;Volume 13:987-1006. [doi: 10.2147/ndt.s130994]
98. 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] [Medline: 27745683]
99. Lustria MLA, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A Meta-Analysis of Web-Delivered Tailored Health Behavior Change Interventions. Journal of Health Communication 2013 Sep;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727]
100. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: A Mechanism to Deliver More Accessible, More Effective Mental Health Care. Clin. Psychol. Psychother 2013 Aug 05;21(5):427-436. [doi: 10.1002/cpp.1855]

Abbreviations

CBITS: Center for Behavioral Intervention Technologies
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders–fourth version
F2F: face-to-face

Edited by J Torous; submitted 21.05.20; peer-reviewed by B Buck, B Johnson; comments to author 12.06.20; revised version received 13.11.20; accepted 26.01.21; published 12.04.21.

Please cite as:
Jonathan GK, Dopke CA, Michaels T, Bank A, Martin CR, Adhikari K, Krakauer RL, Ryan C, McBride A, Babington P, Frauenhofer E, Silver J, Capra C, Simon M, Begale M, Mohr DC, Goulding EH
A Smartphone-Based Self-management Intervention for Bipolar Disorder (LiveWell): User-Centered Development Approach
JMIR Ment Health 2021;8(4):e20424
URL: https://mental.jmir.org/2021/4/e20424
doi:10.2196/20424
PMID:33843607

Original Paper

Measuring and Quantifying Collateral Information in Psychiatry: Development and Preliminary Validation of the McLean Collateral Information and Clinical Actionability Scale

Praise Owoyemi[1*], BA; Sarah Salcone[2*], BA; Christopher King[3*], EdM; Heejung Julie Kim[4*], BA; Kerry James Ressler[5,6*], MD, PhD; Ipsit Vihang Vahia[4,6*], MD

[1]Department of Psychology, Univerity of California, Los Angeles, Los Angeles, CA, United States
[2]Department of Psychology, University of South Alabama, Mobile, AL, United States
[3]Department of Psychology, Emory University, Atlanta, GA, United States
[4]Division of Geriatric Psychiatry, McLean Hospital, Belmont, MA, United States
[5]Division of Depression and Anxiety Disorders, McLean Hospital, Belmont, MA, United States
[6]Department of Psychiatry, Harvard Medical School, Boston, MA, United States
[*]all authors contributed equally

Corresponding Author:
Ipsit Vihang Vahia, MD
Division of Geriatric Psychiatry
McLean Hospital
115 Mill St.
Mail stop 234
Belmont, MA, 02478
United States
Phone: 1 617 855 3291
Email: ivahia@mclean.harvard.edu

Abstract

Background: The review of collateral information is an essential component of patient care. Although this is standard practice, minimal research has been done to quantify collateral information collection and to understand how collateral information translates to clinical decision making. To address this, we developed and piloted a novel measure (the McLean Collateral Information and Clinical Actionability Scale [M-CICAS]) to evaluate the types and number of collateral sources viewed and the resulting actions made in a psychiatric setting.

Objective: This study aims to test the feasibility of the M-CICAS, validate this measure against clinician notes via medical records, and evaluate whether reviewing a higher volume of collateral sources is associated with more clinical actions taken.

Methods: For the M-CICAS, we developed a three-part instrument, focusing on measuring collateral sources reviewed, clinical actions taken, and shared decision making between the clinician and patient. To determine feasibility and preliminary validity, we piloted this measure among clinicians providing psychotherapy at McLean Hospital. These clinicians (n=7) completed the M-CICAS after individual clinical sessions with 89 distinct patient encounters. Scales were completed by clinicians only once for each patient during routine follow-up visits. After clinicians completed these scales, researchers conducted chart reviews by completing the M-CICAS using only the clinician's corresponding note from that session. For the analyses, we generated summary scores for the number of collateral sources and clinical actions for each encounter. We examined Pearson correlation coefficients to assess interrater reliability between clinicians and chart reviewers, and simple univariate regression modeling followed by multilevel mixed effects regression modeling to test the relationship between collateral information accessed and clinical actions taken.

Results: The study staff had high interrater reliability on the M-CICAS for the sources reviewed (r=0.98; P<.001) and actions taken (r=0.97; P<.001). Clinician and study staff ratings were moderately correlated and statistically significant on the M-CICAS summary scores for the sources viewed (r=0.24, P=.02 and r=0.25, P=.02, respectively). Univariate regression modeling with a two-tailed test demonstrated a significant association between collateral sources and clinical actions taken when clinicians

completed the M-CICAS (β=.27; t_{87}=2.47; *P*=.02). The multilevel fixed slopes random intercepts model confirmed a significant association even when accounting for clinician differences (β=.23; t_{57}=2.13; *P*=.04).

Conclusions: This pilot study established the feasibility and preliminary validity of the M-CICAS in assessing collateral sources and clinical decision making in psychiatry. This study also indicated that reviewing more collateral sources may lead to an increased number of clinical actions following a session.

(JMIR Ment Health 2021;8(4):e25050) doi:10.2196/25050

KEYWORDS

electronic media; psychotherapy; text message; electronic mail; collateral information; telecommunication; communications media; digital

Introduction

Background

Reviewing collateral information is an established practice in providing effective and targeted clinical care, particularly in psychiatry. Collateral information provides clinicians with critical information that they may not otherwise be able to obtain from patients' self-reports. Although the importance of gathering collateral information is recognized [1,2], there has been little research of how it may be most effectively gathered, what types of information may be most informative, and how it impacts the clinical decision-making process. Although the gathering of collateral information is considered a part of routine care, this process merits closer study at this time, particularly with an exponential growth in available data from smart technology, wearable devices, and other sensors [3]. It is now feasible to use digital data to measure a broad range of neuropsychiatric symptoms, including depression, anxiety, insomnia, and apathy [4,5]. Similarly, with so much communication between individuals now happening digitally, there may be objective documentation of conversations that can be accessed as part of clinical care [6]. Thus, there is a need for a greater understanding of the types of information that are most relevant and impactful to care and clinical outcomes. This is especially relevant to psychiatry, where augmenting clinical assessment with digital and other data is beginning to impact care [7]. So far, no approach exists to quantify the types of information accessed by a mental health clinician during a typical visit. In addition to the electronic health record (EHR) review, there is no standardized approach or instrument to quantify or assess the full range of clinical decisions that may have been made during a typical clinical visit. EHR review has been shown to be an inadequate and inefficient approach for this because there is often no appropriate documentation of the entire decision-making process, and available information may be challenging to access systematically [8,9].

Objectives

To further examine this issue, we developed an instrument that can measure the number of collateral information sources (CISs) that a clinician may have accessed during an individual session and the total number of clinical actions taken. In this study, we sought to pilot this instrument while also assessing how the nature and volume of collateral information (including digital data) collected may impact clinical decision making. Thus, this study has 2 primary aims. The first aim is to conduct initial feasibility testing of a new measurement tool, the McLean Collateral Information and Clinical Actionability Scale (M-CICAS), and establish both interrater reliability and preliminary validity of this measure against patients' medical records. Our second aim is to test the hypothesis that accessing a greater amount of collateral information would be associated with a higher number of clinical actions taken by the participating clinician.

Methods

Developing the Measure

In developing our survey instrument, we sought to create a measure that could aggregate the number of sources of collateral information that the clinician accessed over the course of the session and determine which aspects of the clinical history the collateral information contributed to. Our approach was modeled on existing literature documenting the development of measures that were based on the aggregation of clinical actions [10,11]. We also aimed to quantify the number of clinical actions taken by the clinician during that session and determine how this information impacted clinical decision making as well as communication between the clinician and patient. To develop the items in the questionnaire, we adopted a consensus-based approach. As an initial step, the study principal investigator (IVV), in consultation with co-investigators, categorized the different types of collateral sources that may be accessed during clinical assessment and the clinical domains that may be impacted through the review of collateral information. On this basis, we selected the following 5 domains: (1) current clinical history, (2) past clinical history, (3) family history, (4) current functioning, and (5) current psychosocial status. Next, we used a similar consensus-based approach to list various clinical actions taken at the end of the session. We also consulted with clinicians who practiced in specialty psychiatry clinics (eg, geriatrics, child and adolescent, or substance use) to generate a more representative set of options. In addition, we requested input from peers at the University of Pennsylvania and Johns Hopkins University (listed in the Acknowledgments section) who are engaged in similar ongoing research.

Final Survey Measure

The final survey consisted of 12 questions, divided into 3 sections (Multimedia Appendix 1). The first section asked clinicians what CISs they reviewed as part of the clinical session and then provided 11 concrete options as well as a write-in option for other sources. The 11 CIS options offered are as

follows: review of medical records, labs, imaging, patient's digital information, talk to mental health provider, talk to non–mental health provider, talk to family/caregiver, talk to nonfamily significant other, talk to patient's school, talk to government agency, and patient-reported outcome measures. Checking off any of these options or the *other* option would prompt a question regarding whether the use of the checked-off collateral source provided additional information about the patient's past or current clinical history, family history, functioning, or psychosocial status.

The second section of the survey sought to establish the clinical actions that were taken following the clinical session. Clinicians were asked whether they changed their treatment plan, adjusted the intensity of care, or took additional clinical actions. The additional clinical actions included changes regarding medications and/or psychotherapy modalities, calls, recommendations, referrals, clinical requests, and screenings. Finally, the third section of the survey examined the shared decision-making process between the clinician and the patient. Clinicians were asked whether they discussed alternatives, risks, side effects, and/or benefits of any of the treatment changes with their patients. This section was created to establish whether clinical actions are actually shared decisions between a clinician and patient, as there may be variability between practices, clinicians, and theoretical orientations. At the end of the survey, a general question was posed, asking clinicians to rate the extent to which they felt that accessing a patient's electronic data might impact therapy outcomes. We established the reliability and validity of both parts distinctly. In the future, we anticipate that these can potentially serve as stand-alone instruments.

Establishing Reliability

Our measure represents an aggregation of information sources accessed and clinical actions taken for a given session, and each score is specific to that session only. Thus, to establish reliability, we focused primarily on establishing interrater reliability using a common source of information (ie, EHR). We determined that establishing test-retest reliability would not be feasible, given the nature of this scale, because it was designed for cross-sectional assessment only.

Demonstrating Validity

To demonstrate the validity of the M-CICAS, we elected to use the EHR as the *ground truth*. A primary driver for developing this measure was to quantify the sources of information reviewed and the actions taken during a session. We recognize that this approach may not be the most suitable because of multiple issues with the process of EHR documentation. The format of documentation in the EHR is nonstandardized, and different clinicians may apply different levels of detail. However, we determined by consensus that this may be the closest we can get to an objective standard by which to demonstrate the preliminary validity of our measure.

Study Participants and Procedures

Description of Clinicians

We initially recruited clinicians at McLean Hospital in Belmont, Massachusetts, who provided ambulatory care to serve as study participants. The eligibility criteria included credentials to provide psychopharmacology or psychotherapy at McLean Hospital and licensed by the State of Massachusetts; actively practicing either adjunctive or stand-alone, evidence-based psychotherapy; and fluent in English. The study staff reviewed the procedures with eligible clinicians, and participation was voluntary. The study procedures were active, and data were collected over a period of 6 months. The participating clinicians completed the M-CICAS after an individual treatment session. Clinicians were asked to complete the measure after as many individual sessions as were feasible in their regular clinical schedule. Scales were completed by clinicians at follow-up visits with patients (ie, not at the intake or baseline assessments). For this study, a treatment session was defined as a single outpatient appointment providing evidence-based psychotherapy and/or pharmacotherapy. These outpatient sessions could also include family/caregivers/partners in the session as long as the patient was present.

After clinicians completed the M-CICAS, 2 members of the study staff (PO and SS) independently conducted a chart review using the clinical notes recorded by the participating clinician. The staff completed the M-CICAS using only the EHR note for the same encounter for which the clinician had completed the M-CICAS as the only source of information. Thus, we used 2 independently rated EHR-based versions of the M-CICAS to establish the reliability of the measure. Before completing the EHR-based data collection, the staff reviewed how to extract collateral information recorded in the medical record progress note before completing the chart reviews. To reduce potential bias, a separate study staff member entered clinician data, and the chart reviewers were blinded to the clinician data.

Testing Associations Between Collateral Information and Clinical Actions

As described earlier, this study has 2 primary goals. The first aim was to conduct feasibility testing of the M-CICAS, and the second was to assess whether there is an association between the amount of collateral information accessed and the number of actions taken by a clinician. For the 89 clinical encounters measured as part of this study, we assessed the associations between the number of data sources reviewed and clinical actions taken.

Analytic Plan

A set of summary scores was generated for the number of collateral sources used in each clinical encounter and the total number of clinical actions taken following each clinical session. We then tabulated the clinician's demographic data. Pearson correlations were used to determine the interrater reliability between clinicians and chart reviewers. To test the association between collateral sources reviewed and the number of clinical actions taken after a session, we first implemented simple univariate regression modeling without accounting for between-clinician differences. Given that subgroups of participants were nested within individual study clinicians and given that the heterogeneity of study clinicians could plausibly have an effect on the association between collateral sources reviewed and clinical actions taken, we also implemented multilevel mixed effects regression modeling. Simple univariate

regression modeling was first applied to test the relationship between collateral sources accessed and the number of clinical actions taken. Next, multilevel mixed effects regression modeling evaluated whether significant variance in clinical actions taken could be attributed to interclinician differences. This allowed us to investigate whether interclinician differences explained significant variance in clinical actions taken and whether, when accounting for interclinician differences, there remained significant associations between collateral sources reviewed and the number of clinical actions taken. All statistical analyses were conducted using the statistical package R version 4.0.2 (R Foundation).

Study data were collected and managed using REDCap (Research Electronic Data Capture) electronic data capture tools hosted at the McLean Hospital [12]. REDCap is a secure, web-based application designed to support data capture for research studies, providing (1) an intuitive interface for validated data entry, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for importing data from external sources.

Results

Clinician Demographics

Table 1 shows the breakdown of clinician demographic information. On average, clinicians reviewed 1.7 collateral sources per clinical appointment and conducted 1.36 clinical actions following each appointment. There was variability in the amount of survey data that each clinician completed. The range of the number of surveys completed by clinicians was broad, with one clinician completing surveys for 4 clinical sessions and another clinician completing surveys for 25 (mean 12.7, SD 8.3) clinical sessions. Overall, 3 clinicians worked primarily in geriatric settings (representing 42 encounters). The remaining 3 clinicians worked in general adult clinics (42 encounters) and 1 clinician in a child and adolescent clinic (5 encounters). Although we did not measure how long it took clinicians to complete the M-CICAS for each patient, based on the subjective impressions of clinicians 1 and 2, it was between 2 and 3 minutes per patient.

Table 1. Clinician demographics.

Clinician	Patient encounters, n (%)	Age (years)	Gender	Race	Clinical experience (years)	Professional degree	Therapeutic approach or approaches	Collateral sources viewed per patient, mean (SD)	Collateral actions taken per patient, mean (SD)
001	17 (19)	38	Female	White	8	APRN[a]	Cognitive behavioral therapy; supportive psychotherapy; mentalization and mindfulness-based therapies	1.35 (1.06)	1.71 (0.99)
002	21 (24)	41	Male	Asian	15	MD[b]	Psychodynamic; expressive therapy; supportive psychotherapy	2.14 (0.91)	1.67 (1.49)
003	25 (28)	30	Male	Asian	5	MD	Psychodynamic; mentalization and mindfulness-based therapies	1.88 (0.60)	0.76 (0.66)
004	4 (4)	69	Male	White	41	LMHC[c]	Cognitive behavioral therapy; acceptance and commitment therapy; supportive psychotherapy	0.25 (0.50)	0.5 (0.58)
005	9 (10)	46	Male	White	10	PsyD[d]	Cognitive behavioral therapy	1.78 (0.97)	1.44 (0.88)
006	8 (9)	65	Male	South Asian/Indian American	30	MD	Cognitive behavioral therapy; psychodynamic; supportive psychotherapy	0.5 (0.53)	0.88 (0.83)
007	5 (6)	31	Female	White	6	PhD[e]	Cognitive behavioral therapy; exposure therapy; dialectical behavioral therapy	4 (0.0)	2.6 (1.82)

[a]APRN: advanced practice registered nurse.

[b]MD: Doctor of Medicine.

[c]LMHC: licensed mental health counselor.

[d]PsyD: Doctor of Psychology.

[e]PhD: Doctor of Philosophy.

Collateral Sources Viewed

Medical records were the most reviewed collateral sources of information (Table 2). Of the 89 clinical total encounters, 62 (70%) of the clinical sessions involved the clinician accessing medical records, distantly followed by 27 (30%) sessions reviewing information from another mental health provider. The remaining collateral source categories were accessed in less than 24% (21/89) of clinical visits. In 12% (11/89) of clinical appointments, the clinician reported not reviewing the collateral information.

Table 2. Type of collateral source reviewed (N=89).

Collateral source type	Percentage of surveys indicating review, n (%)
Medical records	70 (62)
Talk to mental health provider	30 (29)
Labs	24 (21)
Talk to family or caregiver	22 (20)
Patient reported outcome measures	11 (10)
Talk to patient's school	6 (5)
Other	4 (4)
Imaging	3 (3)
Patient's digital Information	2 (2)
Talk to non–mental health provider	2 (2)
Talk to nonfamily significant other	0 (0)

Clinical Domains Impacted

Of the 89 total clinical encounters, 78 (88%) included a review of collateral information. Of the 78 visits where collateral information was reviewed, clinicians most frequently gained insight about a patient's current clinical or mental status as a result of the review, as evidenced in 78% (61/78) of the visits. In 38% (30/78) visits, clinicians learned new information about a patient's functioning through collateral review; in 33% (26/78) of visits, clinicians gained knowledge about their patient's past clinical history; and in 31% (24/78) of visits, clinicians learned more about the patient's psychosocial status. Finally, in 10% (8/78) of visits, clinicians gathered new information about the patient's family history.

Breakdown of Clinical Actions Taken

Clinicians reported adjusting patient medication after a session in 34% (30/89) surveys. Notably, clinicians wrote a separate action in the *other* category in 30% (27/89) of the responses. The remaining listed actions on the survey did not exceed 18% (16/89) of affirmative responses.

Correlations Between Clinician Self-Report and Independent Staff Reviewers

Table 3 shows the correlations between ratings on the 2 sections of the M-CICAS by the 2 study raters and clinicians. Of note, although clinicians seemed to require only 2 to 3 minutes per patient to complete the M-CICAS, rater 1 reported requiring a mean of 3 minutes and 40 seconds (SD 1 min and 54 s) and rater 2 reported a mean of 3 minutes and 42 seconds (SD 1 min and 55 s). There was a range of 37 seconds for the shortest review to 9 minutes and 39 seconds for the longest per patient to review the EHR note and complete the scale for each visit. We found high interrater reliability between the study staff, both of whom independently completed their respective ratings and were blinded to clinician ratings (r=0.98, P<.001 between raters for sources viewed; r=0.97, P<.001 between raters for clinical actions taken). Comparisons between clinician ratings of CISs viewed (based on their self-report) and staff ratings (based on EHR review of the same visit) achieved moderate effect sizes and were also statistically significant (r=0.24, P=.02 and r=0.25, P=.02, respectively, between raters' and clinicians' ratings of sources viewed). However, the same comparisons on the clinical action subscale were not significant.

Table 3. Correlation matrix of study variables with significance level.[a]

Study Variable	1. Clinician source viewed	2. Clinician actions taken	3. Rater 1 sources viewed	4. Rater 1 actions taken	5. Rater 2 sources viewed	6. Rater 2 actions taken
1. Clinician sources viewed						
r	—[b]					
P value	—					
2. Clinician actions taken						
r	0.26	—				
P value	.02	—				
3. Rater 1 sources viewed						
r	0.24	0.09	—			
P value	.02	.41	—			
4. Rater 1 actions taken						
r	0.06	0.11	0.32	—		
P value	.55	.29	.002	—		
5. Rater 2 sources viewed						
r	0.25	0.07	0.98	0.36	—	
P value	.02	.53	<.001	<.001	—	
6. Rater 2 actions taken						
r	0.06	0.13	0.30	0.97	0.34	—
P value	.58	.24	.004	<.001	.001	—

[a]Relationship between number of information sources viewed and clinical actions taken.

[b]Not applicable.

Univariate regression modeling with a two-tailed test, not accounting for the clinician group, revealed a significant association between collateral sources reviewed and clinical actions taken when self-evaluated by clinicians (β=.27; t_{87}=2.47; *P*=.02). To investigate whether this association was significant when accounting for clinicians, we first implemented multilevel random slopes and random intercept models. Analysis of variance tests indicated no significant differences in slopes between the clinician groups (*P*=.11). Consequently, we opted for a multilevel fixed slopes random intercepts model. Even when intercepts were allowed to vary by clinician within the model, there was a significant association between self-evaluated chart sources and clinician actions (Table 4).

Table 4. Clinician self-evaluated chart sources predicting clinician self-evaluated actions taken.

Predictors	Estimates	95% CI	*P* value
Intercept	0.04	−0.28 to 0.35	.83
Clinician sources viewed	0.24	0.02 to 0.45	.03
Random effects			
Variance	0.85	N/A[a]	N/A
Between clinician variance	0.10	N/A	N/A
Intraclass correlation coefficient	0.11	N/A	N/A
$N_{clinician}$	7	N/A	N/A
Observations	89	N/A	N/A
Marginal R^2	0.056	N/A	N/A
Conditional R^2	0.156	N/A	N/A

[a]N/A: not applicable.

Discussion

Principal Findings

The primary goal of this study is to develop and conduct initial feasibility testing of a new two-part measure to quantify the number and types of sources of collateral information accessed by clinicians during a given session and the number of clinical actions taken during that session. We developed this measure based on input from several clinicians across 3 academic departments of psychiatry at McLean Hospital, the University of Pennsylvania, and Johns Hopkins University. We were able to demonstrate the feasibility of using this measure with 7 clinicians from a range of professional backgrounds across 89 patient encounters. Using chart review as the gold standard, we noted that the measure demonstrated acceptable validity (as measured by comparing clinician ratings during the session to staff rating based on chart review). We also noted that 2 independent staff raters had highly correlated scores on both sections of the M-CICAS when the staff raters scored it based on chart review, indicating acceptable reliability. Although the nature of the measure does not facilitate the demonstration of test-retest reliability or true construct validity, our approach is consistent with prior studies on similar measures [11].

We also noted a significant relationship between the number of sources of information reviewed and the number of actions taken (or treatment changes made) by clinicians. Although it is not clear whether a greater number of clinical actions leads to better clinical outcomes, this finding does point to the impact of collateral information on care.

Nonetheless, the implications of this finding are broader than those of this specific study. At a time when extraordinary amounts of information through digital sources are available to clinicians, a major undetermined question is whether access to this information may actually improve care [13]. Our findings provide an early signal suggesting that using collateral data from more diverse sources may have a positive impact on clinical care.

In a sense, our finding is consistent with a vast body of literature that suggests measurement-based care can improve outcomes [14,15]. However, we focused primarily on collateral information rather than quantifying symptom improvement. The M-CICAS may provide a way for researchers to focus on digital phenotyping and generate markers from the sources of digital data to differentiate which types of additional digital information are most relevant and impactful in clinical care. Thus, this scale may play a role in bridging the translation gap from proof-of-concept research on digital health into scaled implementation.

There are a number of limitations to our study. As this measure is an aggregation of the number of actions taken, the concept of construct validity is not applicable. Furthermore, because each scoring of the measure is applicable only to a single session, it was not possible to test true reliability except with the help of a chart review. Thus, our only gold standard to measure both validity and reliability involved reviewing progress notes of the session for which the clinicians completed the M-CICAS. This introduces the possibility of both recall bias and confounds from a lack of standardization in clinical documentation. Other methods to measure validity, such as audio recording or video recording of sessions or direct observation, may provide a higher level of objectivity; however, in this study, the logistical burden of these approaches was not feasible. In addition, as clinicians selected which patients they completed this survey with, the patient sample may not be representative. The relatively small sample size may also limit the generalizability of these findings. Finally, there was a broad range of completed surveys between individual clinicians, which may have introduced bias. Although we believe that our approach to analysis mitigates this effect, the impact of stylistic variations in clinical practice may remain. Nonetheless, we believe that as this measure is largely an aggregation of distinct actions taken by clinicians during a visit, the burden of establishing validity and reliability is lower because the measure is not intended to serve as a measure of abstract behavioral constructs.

Conclusions

In summary, this study demonstrates the feasibility and utility and establishes baseline psychometric properties of the M-CICAS—a new measure that can quantify collateral information and clinical actionability in psychiatric care. Both these entities have been an integral element of clinical care for over a century, but their systematic measurement has not been a focus of research. This study also indicates that reviewing more sources of clinical information may be associated with greater amounts of clinical actions taken at a given session. When the availability of vast amounts of digital information places new burdens on clinicians, this measure may provide a way to determine what types of digital data are most relevant and impactful in patient care. As such, there has been very sparse research assessing how collateral information is collected and used. Our measure and this study represent only an initial step toward quantifying the collateral information used in clinical care. We intend for our approach to serve as a framework and expect that it may evolve to reflect new insights gained with broader application in more studies. We also anticipate that researchers and clinicians may adopt this scale to suit specific studies or clinical quality improvement projects. Our study also points to the possibility that reviewing more sources of clinical information may be associated with greater amounts of clinical actions taken at a given session, although this finding must be replicated in clinician and patient samples that are more standardized. Although this is a preliminary study that merits replication with larger representative samples, we believe that our approach may lay the foundation for a line of research that will facilitate more systematic translation of digital tools into psychiatric patient care.

Acknowledgments

The authors acknowledge the contributions of Courtney Beard, Rose May, Aniqa Rahman, Miranda Skurla, Hannah Heintz, and David Harper in the process of developing the M-CICAS. The authors thank Sophia Shaw for her contributions to data collection and management. The authors also thank Meg Chisholm, Peter Zandi, Leslie Miller, and Tenzin Lhaksampa from Johns Hopkins University and Raina Merchant and Lauren Southwick from the University of Pennsylvania for their feedback on the initial drafts of the measure.

Conflicts of Interest

IVV receives an honorarium from the American Journal of Geriatric Psychiatric for an editorial role. KJR has received consulting income from Alkermes, Bioxcel and Takeda, research support from National Institutes of Health, Genomind, Alto Neuroscience, and Brainsway, and he is on scientific advisory boards for Janssen and Verily, all of which are unrelated to the present work.

Multimedia Appendix 1
McLean Collateral Information and Clinical Actionability Scale (M-CICAS).
[DOCX File , 18 KB - mental_v8i4e25050_app1.docx]

References

1. Remick RA, Sadovnick AD, Gimbarzevsky B, Lam RW, Zis AP, Huggins MJ. Obtaining a Family Psychiatric History: Is it Worth the Effort? Canadian Journal of Psychiatry 1993;38(9):590-594 [FREE Full text] [doi: 10.1177/070674379303800904]
2. Houghtalen RP, McIntyre JS. Psychiatric Interview, History and Mental Status Examination of the Adult Patient. In: Sadock BJ, Sadock VA, Ruiz P, editors. Kaplan & Sadock's Comprehensive Textbook of Psychiatry. Tenth edition. Philadelphia, PA: Wolters Kluwer; 2017.
3. Torous J, Staples P, Onnela JP. Realizing the Potential of Mobile Mental Health: New Methods for New Data in Psychiatry. Current Psychiatry Reports 2015;17(8):602 [FREE Full text] [doi: 10.1007/s11920-015-0602-0]
4. Onnela JP, Rauch SL. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. Neuropsychopharmacology 2016;41(7):1691-1696 [FREE Full text] [doi: 10.1038/npp.2016.7]
5. Bader CS, Skurla M, Vahia IV. Technology in the Assessment, Treatment, and Management of Depression: Harvard Review of Psychiatry. Harvard Review of Psychiatry 2020;28(1):60-66 [FREE Full text] [doi: 10.1097/hrp.0000000000000235]
6. Hobbs KW, Monette PJ, Owoyemi P, Beard C, Rauch SL, Ressler KJ, et al. Incorporating Information From Electronic and Social Media Into Psychiatric and Psychotherapeutic Patient Care: Survey Among Clinicians. Journal of Medical Internet Research 2019;21(7):e13218 [FREE Full text] [doi: 10.2196/13218]
7. Bauer R, Glenn T, Monteith S, Whybrow PC, Bauer M. Survey of psychiatrist use of digital technology in clinical practice. International Journal of Bipolar Disorders 2020;8(1):29 [FREE Full text] [doi: 10.1186/s40345-020-00194-1]
8. Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L. Promises and pitfalls of electronic health record analysis. Diabetologia 2018;61(6):1241-1248 [FREE Full text] [doi: 10.1007/s00125-017-4518-6]
9. Yin AL, Gheissari P, Lin IW, Sobolev M, Pollack JP, Cole C, et al. Role of Technology in Self-Assessment and Feedback Among Hospitalist Physicians: Semistructured Interviews and Thematic Analysis. Journal of Medical Internet Research 2020;22(11):e23299 [FREE Full text] [doi: 10.2196/23299]
10. Cacciola JS, Alterman AI, Lynch KG, Martin JM, Beauchamp ML, McLellan AT. Initial reliability and validity studies of the revised Treatment Services Review (TSR-6). Drug and Alcohol Dependence 2008 Jan 01;92(1-3):37-47 [FREE Full text] [doi: 10.1016/j.drugalcdep.2007.06.004]
11. McLellan AT, Alterman AI, Cacciola J, Metzger D, O'brien CP. A new measure of substance abuse treatment. Initial studies of the treatment services review. The Journal of Nervous and Mental Disease 1992;180(2):101-110 [FREE Full text] [doi: 10.1097/00005053-199202000-00007]
12. 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. Journal of Biomedical Informatics 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010]
13. Gillan CM, Whelan R. What big data can do for treatment in psychiatry. Current Opinion in Behavioral Sciences 2017;18:34-42 [FREE Full text] [doi: 10.1016/j.cobeha.2017.07.003]
14. Trivedi MH, Papakostas GI, Jackson WC, Rafeyan R. Implementing Measurement-Based Care to Determine and Treat Inadequate Response. Journal of Clinical Psychiatry 2020;81(3) [FREE Full text] [doi: 10.4088/jcp.ot19037br1]
15. Turvey CL, Lindsay JA, Chasco EE, Klein DM, Fuhrmeister LA, Dindo LN. Current Practices in Electronic Capture of Patient-Reported Outcomes for Measurement-Based Care and the Use of Patient Portals to Support Behavioral Health. Psychiatric Clinics of North America 2019;42(4):635-647 [FREE Full text] [doi: 10.1016/j.psc.2019.08.006]

Abbreviations

CIS: collateral information source
EHR: electronic health record
M-CICAS: McLean Collateral Information and Clinical Actionability Scale
REDCap: Research Electronic Data Capture

Edited by J Torous; submitted 20.10.20; peer-reviewed by E Lee, J Torous; comments to author 03.12.20; revised version received 22.12.20; accepted 14.01.21; published 14.04.21.

Please cite as:
Owoyemi P, Salcone S, King C, Kim HJ, Ressler KJ, Vahia IV
Measuring and Quantifying Collateral Information in Psychiatry: Development and Preliminary Validation of the McLean Collateral Information and Clinical Actionability Scale
JMIR Ment Health 2021;8(4):e25050
URL: https://mental.jmir.org/2021/4/e25050
doi:10.2196/25050
PMID:33851928

Original Paper

Youth and Provider Perspectives on Behavior-Tracking Mobile Apps: Qualitative Analysis

Courtney C Armstrong[1], BS; Erica J Odukoya[2,3], MPH; Keerthi Sundaramurthy[1], MAT; Sabrina M Darrow[3], PhD

[1]Department of Psychology, University of California, Berkeley, Berkeley, CA, United States

[2]University of Michigan Medical School, Ann Arbor, MI, United States

[3]Department of Psychiartry, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Courtney C Armstrong, BS
Department of Psychology
University of California, Berkeley
2121 Berkeley Way
University of California
Berkeley, CA, 94720
United States
Phone: 1 9178417015
Email: courtney.armstrong@berkeley.edu

Abstract

Background: Mobile health apps stand as one possible means of improving evidence-based mental health interventions for youth. However, a better understanding of youth and provider perspectives is necessary to support widespread implementation.

Objective: The objective of this research was to explore both youth and provider perspectives on using mobile apps to enhance evidence-based clinical care, with an emphasis on gathering perspectives on behavior-tracking apps.

Methods: Inductive qualitative analysis was conducted on data obtained from semistructured interviews held with 10 youths who received psychotherapy and 12 mental health care providers who conducted therapy with youths aged 13-26 years. Interviews were independently coded by multiple coders and consensus meetings were held to establish reliability.

Results: During the interviews, the youths and providers broadly agreed on the benefits of behavior tracking and believed that tracking via app could be more enjoyable and accessible. Providers and youths also shared similar concerns that negative emotions and user burden could limit app usage. Participants also suggested potential app features that, if implemented, would help meet the clinical needs of providers and support long-term use among youth. Such features included having a pleasant user interface, reminders for clients, and graphical output of data to clients and providers.

Conclusions: Youths and providers explained that the integration of mobile health into psychotherapy has the potential to make treatment, particularly behavior tracking, easy and more accessible. However, both groups had concerns about the increased burden that could be placed on the clients and providers.

(JMIR Ment Health 2021;8(4):e24482) doi:10.2196/24482

KEYWORDS

qualitative; mHealth; mobile phone; behavior monitoring; youth

Introduction

Problems with mental health are common among youth. Prevalence data suggest that 1 in 5 children has experienced symptoms of a mental disorder and 1 in 10 has experienced a serious emotional disturbance [1,2]. Mental health disorders that emerge during adolescence and early adulthood frequently continue into adulthood and contribute significantly to the existing global burden of disease [3,4]. Evidence-based interventions to treat mental illness have been developed and tested [5-7]. However, these interventions are not widely available, and of those who are able to receive treatment, younger clients are less likely to engage [6,8-11]. The integration of mobile apps into evidence-based interventions is one potential pathway to improving engagement and outcomes for adolescents and young adults (hereon, referred to as youth) [12]. As smartphone usage across all age groups has increased, the development and usage of mobile apps for the management of

health has also increased immensely [13], and many mobile health (mHealth) apps are designed specifically for improving mental health [14,15]. In industrialized nations, the majority of youth have access to a smartphone or a tablet [16,17], thereby making apps increasingly accessible [18]. Prior research suggests that adolescents are interested in incorporating mobile apps into their treatment [19], and one study found that a mobile app that was integrated into treatment-as-usual with adults was effective and well received [20]. However, the initial promise of mobile apps has not been readily translated into widespread use. Despite their abundance, few apps designed specifically for youth are evidence-based [21]. Furthermore, previous research suggests that youth struggle to maintain long-term app usage [21,22], which can be a problem if attrition occurs before goals are reached or skills are fully learned. Given these considerations, there is a need for increased focus on user-centered approaches to mHealth [23].

Using mHealth apps with support from a health care provider has the potential to increase long-term engagement for youth [23,24]. Unfortunately, many health care providers are hesitant to integrate mHealth into their practice [25]. There are many factors that influence this, including poor fit to the treatment context or population and a lack of research exploring the perspectives of providers when it comes to integrating mHealth apps into routine care. Given that mental health interventions are unlikely to be adopted into practice or fully implemented if they do not fit the context of care, this is a significant barrier [26,27]. Thus, to achieve widespread implementation of mHealth apps, a better understanding of providers' needs is imperative.

This study uses a qualitative approach to explore youth and mental health provider perspectives on using apps to enhance evidence-based clinical care, with an emphasis on gathering perspectives on behavior-tracking apps. Tracking behaviors, including mood and thoughts, is an effective component to many evidence-based treatments that leads to tangible changes in behavior and improved therapeutic outcomes [28]. Further, it is easily translated into an electronic task [29]. By exploring providers' perspectives, this research highlights how mHealth apps could address the challenges providers face. Additionally, gathering the perspectives of youths who receive treatment will help identify the app features that will engage young clients. In-depth information on these topics may provide insight for future app designers, thereby ultimately improving the odds of successfully implementing an app to be used alongside face-to-face treatment.

Methods

Design

This was a qualitative, individual interview study to gather perspectives on mHealth usage from youth clients and providers. Interviews were conducted as part of an effort to develop a mobile app to be used alongside outpatient therapy for youth. Responses from the interviews were organized into codes and categorized into broader domains. All research activities were approved by the appropriate Institutional Review Board.

Participants

Youth

A convenience sample of youths were recruited via study flyers posted within San Francisco. The inclusion criteria for participation were (1) between the ages of 13 and 26 years, (2) ability to speak English, and (3) participation in at least one session of outpatient psychotherapy. Youths were excluded if they had a visual, hearing, voice, or motor impairment preventing the use of mobile phones. Of the 10 individuals who contacted study personnel, all met the study criteria and were interviewed.

Providers

Recruitment emails were sent to providers working with youth at an academic medical center and to a listserv of community providers. The inclusion criteria for participation were (1) they currently conduct psychotherapy with youths between the ages of 13 and 26 years and (2) ability to speak English. Providers were excluded if they had a visual, hearing, voice, or motor impairment that interfered with use of mHealth apps. Of the 12 providers who contacted the study team, all met the inclusion criteria and were interviewed. The sample included 6 medical center providers and 6 community providers.

Materials

Youth Interview Guide

Youth interviews consisted of 8 open-ended questions, followed by prompts to facilitate further discussion. Youths were asked questions regarding their preferences when selecting mobile apps, previous experiences tracking behavior during psychotherapy, app features that contribute to a positive user experience, and considerations that may facilitate the use of apps for tracking behavior during psychotherapy.

Provider Interview Guide

Provider interviews consisted of 5 open-ended questions, followed by prompts. Providers were asked to reflect on the outcomes of behavior tracking with clients and factors that affected the completion of behavior tracking. Prompts were used to determine the role of these factors in treatment-related decisions. Additionally, providers were asked to discuss current methods of behavior tracking, potential areas of improvement, and barriers to implementing electronic behavior tracking with clients.

Procedure

Prior to the interview, participants completed an electronic consent form and a brief demographics survey using Qualtrics, a web-based survey platform. Telephone interviews were conducted to accommodate community providers; the remainder were conducted in-person at an academic medical center. Interviews took place in private spaces with no nonparticipants present. Repeat interviews were not conducted. Interviews were conducted by 1 of the 4 potential interviewers and were approximately 35 minutes in length (median 34.8 minutes, IQR 17.60 minutes). To begin, trained interviewers explained the purpose of the interview and reminded participants that they would be audio-recorded. Interviewers then audio-recorded and

conducted the interview following the interview guides. Interviewers were asked to take notes during the interview process in order to assist with coding. Participants were emailed a US $30 electronic gift card. Interviews were conducted until no new knowledge was being obtained from the new participants.

Data Analysis

Recordings were imported into the qualitative data analysis software Atlas.ti version 7 for coding and analysis. Two trained coders (EO and KS) reviewed each interview individually using a general inductive approach to analyze the data [30]. During the initial review, each coder listened for thematic content expressed by the participants related to the objectives of this research. Through discussion, the coders identified patterns in the data and created an initial code list to assign to portions of the recordings. During subsequent review, coders assigned codes originating from the previous discussions to quotations, and these quotations were transcribed. Coders met regularly to discuss discrepancies, develop new codes, and revise code definitions. As codes were further refined, interviews were continuously reviewed to adjust coding. Disagreements between the coders were resolved until complete agreement was reached by reviewing transcribed quotations and interview recordings.

Results

Participants in This Study

A total of 22 participants were interviewed for this study. This pool of participants consisted of 10 youths and 12 providers. The demographics of the youths and providers are illustrated separately in Table 1 and Table 2, respectively.

Table 1. Demographical data of the youths in this study (n=10).

Characteristic	Values
Age (years), mean (SD)	18.9 (3.73)
Gender (female), n (%)	9 (90)
Relationship status (single), n (%)	8 (80)
Sexuality, n (%)	
Heterosexual	9 (90)
Bisexual	1 (10)
Homosexual	0 (0)
Other	0 (0)
Race/ethnicity, n (%)	
White/Caucasian	4 (40)
Black/African American	1 (10)
Asian	2 (20)
Native Hawaiian/Pacific Islander	1 (10)
Mixed race	2 (20)
Residential environment, n (%)	
Urban	8 (80)
Suburban	2 (20)
Highest education level, n (%)	
8th grade or less	1 (10)
Some high school	4 (40)
Graduated high school or obtained general education diploma	1 (10)
Graduated 4-year college	4 (40)
Completed graduate or professional school	0 (0)
Highest parental education level, n (%)	
8th grade or less	0 (0)
Some high school	0 (0)
Graduated high school or obtained general education diploma	1 (10)
Graduated 4-year college	4 (40)
Completed graduate or professional school	5 (50)
Currently in therapy, n (%)	8 (80)
Length of time in therapy, n (%)	
<1 year	2 (20)
1-2 years	1 (10)
2-5 years	5 (50)
5+ years	2 (20)
Taking medication for mental health, n (%)	8 (80)
Employment status, n (%)	
Full-time employment	1 (10)
Full-time student	5 (50)
Part-time employment	2 (20)
Employed, full-time student	1 (10)
Unemployed	1 (10)

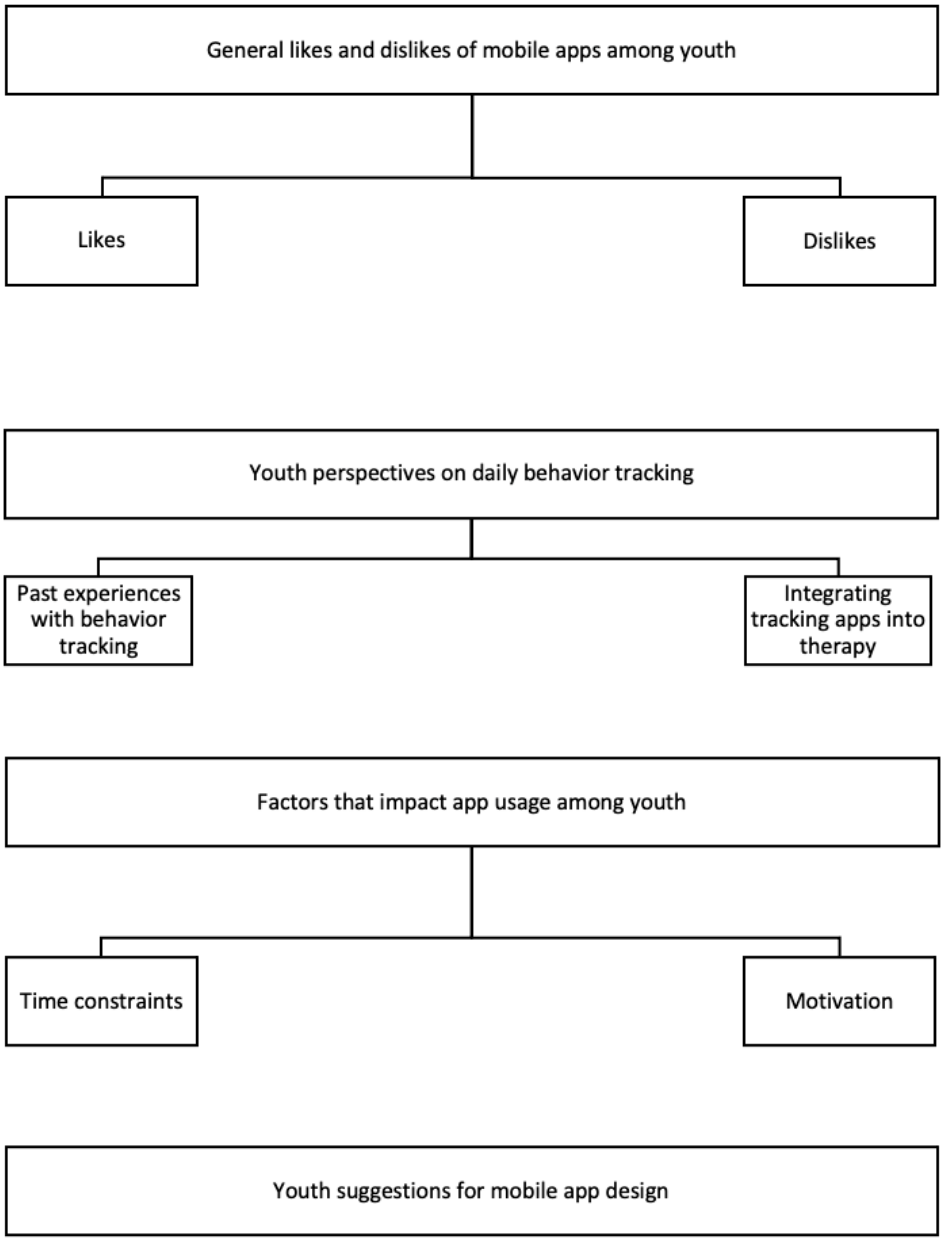
Table 2. Demographical data of the mental health care providers in this study (n=12).

Characteristic	Values
Age (years), mean (SD)	38.42 (5.78)
Years since received degree, mean (SD)	11.08 (3.8)
Proportion of time doing clinical work, mean (SD)	56.08 (30.65)
Proportion of clinical work with adolescents, mean (SD)	36.25 (20.35)
Proportion of clinical work with adults, mean (SD)	49.17 (25.03)
Gender (female), n (%)	8 (67)
Relationship status (single), n (%)	1 (8)
Licensed, n (%)	11 (92)
Employment status, n (%)	
Full-time	10 (83)
Part-time	2 (17)
Race/ethnicity, n (%)	
White/Caucasian	9 (75)
Black/African American	0 (0)
Asian	2 (17)
Native Hawaiian/Pacific Islander	0 (0)
Mixed race	1 (8)
Sexuality, n (%)	
Heterosexual	9 (75)
Bisexual	1 (8)
Homosexual	1 (8)
Other	1 (8)
Profession, n (%)	
Psychologist	6 (50)
Social worker	1 (8)
Licensed marriage and family therapist	5 (42)
Theoretical background, n (%)	
Behavioral	1 (8)
Cognitive behavioral	5 (42)
Family systems	2 (17)
Patient-centered	1 (8)
Positive and self-compassion based	1 (8)
Eclectic	2 (17)

Youth Interviews

Codes from the 10 youth interviews were sorted into 4 broad domains: (1) general likes and dislikes of mobile apps, (2) perspectives on daily behavior tracking, (3) factors that affect app usage, and (4) suggestions for mobile app design. These domains were broken up into subthemes. The domains and relevant subthemes are illustrated in Figure 1.

Figure 1. Summary of youth domains and subthemes.



General Likes and Dislikes of Mobile Apps Among Youth

Likes

All interviewed youths identified the mobile app features that they enjoyed. They identified 4 major positive features. The most commonly discussed feature was user-friendliness, which was defined by a participant as an app that "isn't time-consuming or complicated to learn how to use." [Youth 1] Other participants described enjoying apps that were easy to understand and made entering information a quick and efficient process. Youths also enjoyed receiving encouraging messages from apps and apps that helped them cope with stressors outside of therapy sessions.

Dislikes

All youths described negative experiences and the corresponding app features they disliked. They identified 4 major dislikes. First, the majority of the youths disliked apps that broke down frequently, (eg, issues with freezing, glitches, connectivity, other technical errors) preventing the app's use. Second, a smaller group of participants reported frustration with apps that gave inaccurate or poorly organized information. These apps created more work and increased user burden. An example provided was auto-scrolling, which obstructs access desired to information. Additionally, participants disliked when information was obscured by advertisements. Third, another smaller group of youth disliked poorly formatted apps or apps that frequently changed their formatting, as this increased the effort expended to relearn how to use the app. Fourth, a small minority of participants expressed concerns with the lack of privacy; reasons included disliking any app that would not treat data with privacy and the ease with which others could find their information. Taken together, youth likes and dislikes strongly emphasize the importance of an attractive yet simple

and consistent user interface that allows information to be made readily accessible.

Youth Perspectives on Daily Behavior Tracking

Past Experiences With Behavior Tracking

Among the subset of participants who reported previously tracking behaviors as part of therapy, perspectives were mixed. Participants recognized that tracking behavior brought positive benefits to therapy, including improved quality of discussion with their providers. Additionally, tracking helped the youths set goals, monitor progress, and regulate behavior. However, the process of tracking behavior could be unenjoyable and anxiety-inducing for some participants. Additionally, 1 participant reported getting frustrated with the tracking over time despite positive first impressions. Another participant found tracking annoying but still perceived the benefits. Youths who had a positive or initially positive experience with tracking reported that they enjoyed the process and that tracking improved discussions with their providers. Among youths who disliked tracking behavior, a key commonality was feeling discouraged after entering information into a tracker and struggling to meet goals.

Integrating Tracking Apps Into Therapy

The youths gave mixed responses about using apps to track activities in therapy. Most provided reasons why using an app would be more beneficial than recording activities on paper. The reasons included increased convenience, more accurate tracking, improved privacy, more room for detail, and the ability to visually monitor progress with graphs and charts. A participant also noted that using an app to track activities would save paper. A smaller subset of the youth participants also shared reasons on why tracking activities with an app could be detrimental. They noted that tracking with an app was less personal and that the process of writing fostered a connection to what was being written down.

Factors That Affect App Usage Among Youth

Time Constraints

Most youths reported that they would feel uncomfortable using their phones to track activities in certain situations (eg, at work, school, studying, with friends). One reported that this could be combated by setting time aside at the end of the day to record everything, which they preferred over monitoring activities throughout the day.

Motivation

A few youths noted that their own dislike for behavior tracking could lower personal motivation and make tracking harder, though behavior tracking would be easier if they considered it important. In comparison, a slightly larger subset noted that behavior tracking could be made easier by experiencing positive outcomes, such as providing content for therapy sessions and improved communication with their providers.

Youth Suggestions for Mobile App Design

Youth participants suggested numerous features and considerations for a mobile app that would best suit their needs and facilitate behavior tracking. The most frequently made suggestion was that the app should have an interface that is both pleasing and easy to navigate. Additional features suggested by the youths can be found in Table 3.

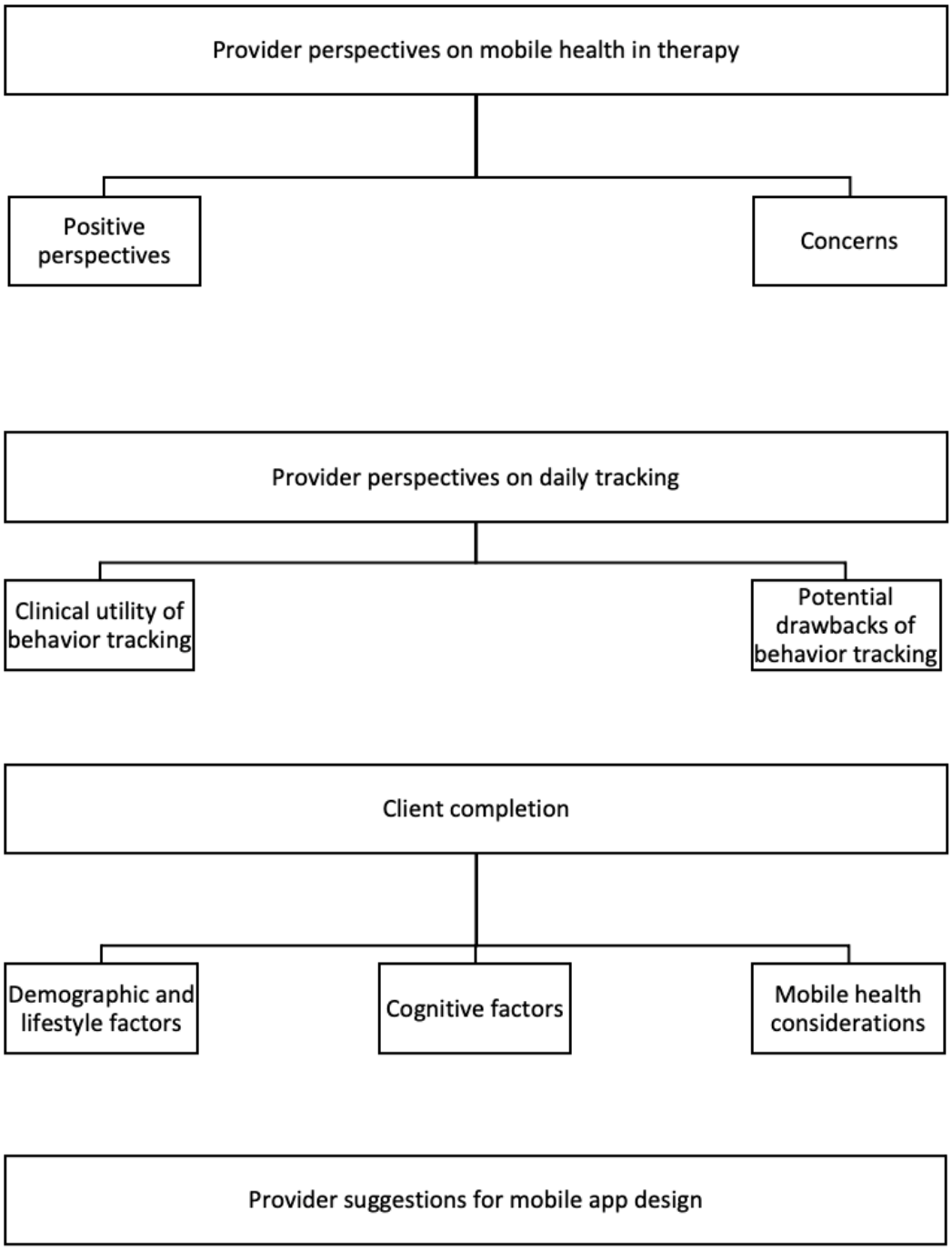
Table 3. Summary of the suggestions provided by the youths for mobile health app features.

Function of feature, feature	Definition	Illustrative quotes
Ease of use		
Reminders	Notifications reminding clients to complete their tracking	*…You'd want the app to send you some sort of notification. Like, 'hey little check in: did this happen? Do you want to log it?* [Youth 9]
Connecting to computer	Usable on multiple technological platforms	*…Something that can connect to my computer would be nice.* [Youth 7]
Improve appeal of use		
Attractive design	Having an aesthetically pleasing and simple layout	*…You should be able to go in and not get frustrated while doing what you need to do.* [Youth 2]
Customization	Ability to personalize activities and data input	*…I like how many options they have for tracking things. It makes it feel a lot more personal.* [Youth 5]
Reinforcement	A feature that rewards activity completion	*…When you meet your goal, everybody wants a little pat on the back.* [Youth 3]
Provide additional information		
Communication with provider	Enabled communication between client and provider outside of sessions	*…It would be great if there could be a messaging system where the therapist asks you questions about things that you say.* [Youth 4]
Output to client	Organize/summarize data for easy review	*…Charts are important because it's a visual way to see the progress you've made or if you're slipping.* [Youth 1]
Additional resources	Clients or providers able to add relevant information to be accessed between sessions	*…It would be great if skills you used could be saved.* [Youth 4]

Provider Interviews

The codes from the 12 provider interviews were categorized into 4 umbrella domains: (1) perspectives on mHealth, (2) perspectives on behavior tracking (3) client compliance, and (4) suggestions for mobile app design. These domains were broken into subthemes. The following sections present qualitative data from the provider interviews describing these subthemes (see Figure 2 for overview).

Figure 2. Summary of provider domains and subthemes.



Provider Perspectives on mHealth in Therapy

Positive Perspectives

The most frequently mentioned benefits to the client included that apps are easy to use and more attractive to younger clients when compared to pencil and paper. Two providers reported that it would be useful for clients to frequently provide updates on their lives in the moment. Nearly all providers discussed ways in which electronic behavior tracking could make their clinical work easier. Primarily, these providers spoke of how behavior-tracking apps could summarize trends in client behavior by using an organized format.

Concerns

Providers reported 6 primary concerns related to the use of electronics in clinical practice. Most frequently, providers reported that behavior-tracking apps can be confusing and that "if [an app is] not very self-explanatory or not very easy to

access, [clients] won't do it probably." [Provider 3] Half of the providers also expressed that they were worried about apps being inaccessible to clients either due to Wi-Fi and connectivity issues or for financial reasons. One provider voiced discomfort with the idea of increasing their clients' use of technology and suggested that writing instead of typing would help clients better process information. Some providers were concerned that incorporating technology into their practice would lead to unnecessary additional work and frustration. They noted that receiving too much information too frequently can be overwhelming and mHealth apps might have problems with security or assume a level of technological understanding that not all providers have. Additionally, the notion that electronic tracking needs to have clinical utility was highlighted, "If the data are just going to a data bank, it's kind of static, just like paper...then maybe it's not necessary for tracking to be tech-based." [Provider 12]

Provider Perspectives on Daily Tracking

Clinical Utility of Behavior Tracking

Providers noted that behavior tracking provided important information about clients' behaviors and can facilitate behavior change by increasing awareness of relevant behaviors, thus benefiting both them and their clients (eg, "It's useful for identifying triggers or identifying thinking patterns" [Provider 4]). Some providers shared that behavior tracking was also useful for improving communication, specifically stating that behavior tracking allowed them to make therapy more personal and helped clients set session agendas.

Potential Drawbacks of Behavior Tracking

Providers reported several drawbacks related to behavior tracking. Although individual responses varied, all were related in some way to client completion. A few providers stated that clients often failed to track activities every day, leading to less accurate information. One provider stated that when behavior tracking was incomplete, valuable session time was spent on problem solving. Additionally, some providers stated that tracking every day can feel burdensome to clients, with one also stating that activities done outside of the session can be tracked incorrectly. This is particularly relevant in the context of family therapy, as providers described how difficult it could be to manage daily behavior tracking for multiple people.

Client Completion

Demographic and Lifestyle Factors

Half of the providers reported that older clients are generally more willing to track behaviors outside of session. This was encapsulated by the statement, "Age does help, and the ability to understand the potential utility of [behavior tracking] as a strategy." [Provider 1] Half of the interviewed providers also reported that clients with a higher socioeconomic status were more likely to complete daily behavior tracking, as "lower socioeconomic status clients are more in survival mode, and that will often trump everything else." [Provider 3] Providers also reported specific factors such as lifestyle, education, and family structure were likely to impact completion. Clients who are able to give more of their time and energy to tracking an activity in daily life are more likely to complete. To quote one provider: "Clients where I have seen it not work...have multiple other life burdens in terms of their hierarchy of needs." [Provider 5] Similarly, having family structure and consistency in the household was thought to increase the likelihood of a client tracking daily behaviors. Clients who are in school, complete daily homework, or complete other daily routines were also suggested as more likely to track behaviors consistently. One provider reported race as a potential factor affecting client completion, although specific racial groups were not mentioned.

Cognitive, Behavioral, and Personality-Related Factors

Cognitive, behavioral, and personality-related traits that providers suggested would decrease the likelihood of clients completing daily tracking included issues with attention, motivation, memory, pathology, shame, and nonspecific cognitive difficulties. Specifically, clients who experienced trouble with attention or memory were thought to be likely to struggle with tracking due to difficulties with completing the behavior of interest. Clients who experience difficulties with motor function or vision could have trouble completing a tracking log, and clients who could not comprehend the activity being tracked would be unlikely to complete it. Clients who experience significant shame, specifically related to the behavior being tracked, might also be unwilling to report instances of the behavior to a provider.

Cognitive, behavioral, and personality-related traits positively affecting completion included clients described as responsible, anxious, or having positive behaviors, and feelings related to therapy generally. A few providers reported that "responsible" clients, and clients who were more "rules-oriented" were more likely to complete therapy-related tasks, particularly behavior tracking. Additionally, those who have more trust in their provider or exhibit more buy-in regarding therapy are more likely to complete assignments. Three clinicians explicitly mentioned that youths who had sought out therapy themselves, rather than having parents seek out treatment on their behalf, are more likely to complete tracking.

mHealth Considerations

Providers reported 8 ways that mHealth apps could be used to improve completion. Specifically, providers said that an app would make it easier for clients to complete the activity. Half of the providers thought an app could be used to set up systems of accountability, ensuring that clients complete the assigned activity. Some reported that apps could remind clients to complete tracking activities, although there was some concern that clients would habituate to reminders over time. Providers also reported that apps could provide examples to improve understanding of tasks, pair assigned activities to more enjoyable activities to increase motivation, make assigned activities seem more relevant to the client, make the tracked activities seem less rigid, and reinforce completion.

Provider Suggestions for Mobile App Design

Providers suggested 25 different types of behaviors that they might want to track with an app (Textbox 1).

Providers also suggested multiple features for a behavior-tracking app that would best suit their needs. The most frequently given suggestion was that an app should clearly summarize client data. Additional features suggested by the providers can be found in Table 4.

Textbox 1. Behaviors to track that were suggested by the providers.

- Physical activity
- Emotional intensity
- Conflict
- Device use
- Substance use
- Sleep and appetite
- Attendance
- Irritability
- Problem behaviors
- Social activities
- Thoughts
- Urges
- Community and citizenship
- Fatigue
- Motivation/concentration
- Relationships
- Relaxation activities
- Therapy homework
- Work/school
- Achievement
- Creativity
- Driving
- Money use
- Sexual risk taking
- Extracurricular activities

Table 4. Summary of the provider suggestions for mobile health app features.

Function of feature, feature	Definition	Illustrative quotes
Ease of use		
Simplicity	Having an intuitive and efficient interface	*…Something super intuitive that could be used in a way that is almost dummy-proof. So even if you're not super tech savvy… it would just be so easy. And quick.* [Provider 6]
Reminders	Notifications reminding clients to complete their tracking	*…A reminder is always helpful… I think for some people, an alarm or a popup, a sound, or whatever would be easy for people to see and do. They can just click it and go right into it…being able to set up when and how it shows up, according to thepatients'preferences.* [Provider 7]
Remembering data	Previous entries remembered to ease data entry	*…It would remember what I'd already entered and add it to the menu, so if I had the same breakfast three days, I could just hit breakfast and it would just put in all three.* [Provider 10]
Data-based suggestions and advice	Tips or information about the tracked behavior	*…Kind-of a help button. Any kind of monitoring form can be confusing… when we ask about your mood scales, there's kind of a definition with each question… it's really clear what information is needed.* [Provider 12]
Improve clinical utility		
Customization	Ability to personalize activities and data input	*…My dream application would be a sophisticated application that allows you to create scales and create the flexibility to make whatever scale I want. So, it'd have a bunch of different ways to track something.* [Provider 11]
Provider notifications	Alerts telling the provider when the client enters data	*…[Something that] gets communicated to me via link or email and it's just very clear and understandable and it doesn't take too much time for me to understand it.* [Provider 1]
Output for providers	Organized summaries of client data such as a table or graph	*…If there were a system… that automatically graphed things for you… So I could log in as their provider and see a graph of progress over time.* [Provider 1]
Parent information	Allow parents to enter information	*…Maybe it would be nice if there was a child feature and a parent feature, or if there was an ability to link between phones so the parent and the child could input data.* [Provider 8]
Improve appeal of tracking		
Visually appealing	Pleasing interface rather than plain text	*…Something that reminds the client to take some time and sit down… I wouldn't want the app to be like opening up the notes section of your iPhone.* [Provider 5]
Reinforcement	A feature that rewards activity completion	*…[There should be] immediate reinforcement on completing it. Technology has built up youth's dependence on instant gratification, so using that in a way that would help them continue to be using it.* [Provider 8]
Options	Multiple-choice options for describing mood, rather than an open-ended format	*…Being able to scroll down and select [from a list of] cognitive distortions.* [Provider 2]
Provide additional information		
Output to clients	Data summaries that are accessible to the client	*…If the app can output the data back to the client in a way that helps them see like, 'oh on average you're getting so much sleep!'… Because that's what monitoring is about. To help us better understand these patterns, and make use of these patterns, to make use of it in therapy and treatment and address it.* [Provider 7]
Crisis resources	Provide crisis resources	*…What would be great is having a page or a tab, or something where there's a list of local crisis resources, so text line, lifeline, rape line, so somehow pulling from crisis stuff.* [Provider 9]
Social support	Provide social support or connect to social media	*…It lets me post to social media, so you get that kind of support too.* [Provider 7]

Discussion

Principal Results

This study explored client and provider perspectives on using mHealth apps to enhance clinical care and identified features that they would want in an app designed for daily behavior tracking during clinical care. This study adds to a growing literature exploring provider perspectives on mHealth and is the first to do so while synthesizing youth and provider perspectives, thus improving our understanding of mHealth's potential to improve behavior tracking's accessibility and usage. Both providers and youths considered behavior tracking to be a beneficial activity that led to more meaningful therapy sessions and behavior change. These findings are consistent with the broader literature showing that tracking behavior helps clients maintain healthy attitudes, counteract negative developments, and improve self-management [31]. Furthermore, supplying providers with additional information regarding a client's behavior can improve treatment outcomes [32]. Reasons for this include the possibility that additional information from clients allows providers to tailor session content based on events that occur outside of therapy [31]. Youths and providers both believed that using an app could be a more enjoyable and accessible way to track behavior. Widespread smartphone usage facilitates behavior tracking in situations where clients cannot access paper-based resources used to supplement many mental health interventions (ie, diary cards, handouts, and workbooks). This could increase the accuracy of behavior tracking, further improving clinical utility. Additionally, apps have the potential to summarize collected data in a format that is easy to understand, making progress-tracking easier and facilitating discussion of progress during treatment.

Despite the recognized benefit of daily tracking, views regarding completion and app usage were mixed. Select youth participants found behavior tracking frustrating, which could impede behavior tracking. Providers also believed that negative emotions could hinder behavior tracking, particularly for those too young to fully appreciate its long-term benefits. Youths and providers were also concerned about the added burden of an app. Youth participants disclosed that social environments (eg, school, work) limit opportunities to track behaviors. Likewise, providers stated that individual pathology or factors related to socioeconomic status might negatively impact clients' completion of behavior tracking. To minimize burden on youth clients, providers and youths suggested that an app should be highly customizable so that only relevant information is tracked. It should be noted that many of these challenges related to burden parallel what clinicians already encounter with paper-based tracking.

This problem of increased burden is not solely applicable to clients and highlights a notable way in which client and provider preferences may conflict. Indeed, half of the youth participants wanted an app that would allow for communication with their providers between sessions. However, providers expressed concern that an app could be overwhelming due to increased information received from clients and insufficient understanding of technology. This is a significant barrier to the integration of mHealth into routine care, as evidence-based interventions are difficult to implement when providers lack sufficient training, resources, and support [33]. To circumvent this issue, app designers should prioritize minimizing burden placed on both providers and clients. Features such as reminders and more immediate reinforcement may be helpful for some clients, but they may not work for others. Additionally, some providers may want to communicate with clients via an app, while others may not. Designing a flexible behavior-tracking app that allows for customizations as well as considering how a platform could be used in the case that paper-based tracking will be preferred by some, may be the best solution for widespread implementation. Youths and providers also had concerns about data security and privacy. The use of passwords and secure login portals are potential solutions suggested by both providers and youth. Prior research highlights additional steps that can be taken to ensure client data are secure [34].

Strengths and Limitations

This study is among the first to compare youth and provider perspectives regarding the integration of mHealth into evidence-based, routine psychotherapy practice. Taking the perspectives of both groups allows future researchers and developers to consider the best possible ways of meeting the needs of both groups without placing inhibitory burdens on either. Regular review of the recorded interviews and interview guides allowed for a thorough examination of participant-driven topics. There are limitations to this research. Because the mean age of the youth participants was 18.9 (SD 3.73) years, their perspectives might not fully reflect the views of younger adolescents. Additionally, the youth participants in this study were 90% (9/10) female. However, this is consistent with that reported previously in treatment-seeking populations [35]. Despite our relatively small sample size, participants shared a variety of challenges to implementing an mHealth app as well as many different behaviors that might be clinically indicated for different individuals to track.

Future Directions

This study highlights provider and client perspectives on the acceptability of using behavior-tracking apps in clinical practice. Additionally, the study discusses design features that would better facilitate widespread implementation of behavior-tracking apps. Specifically, providers and clients alike stressed the need for features such as reinforcement, which may assist with long-term use. Flexibility and simple interface design were also deemed important and could help to minimize the burden placed on providers. Future research should explore whether specific features such as reinforcement and interface design reduce the perceptions of burden and facilitate user engagement. Additionally, it may be beneficial to examine the feasibility of tracking apps across specific modalities and treatment environments such as community mental health clinics and private practice.

Conclusions

This study shows that mHealth has the potential to improve daily behavior tracking for youth recipients and providers of mental health services. First, this study highlights that behavior

tracking is generally acceptable to both youths and providers, largely due to the increased information shared between clients and providers. Second, it explores ways in which mHealth can make behavior tracking more accessible and enjoyable for both clients and providers. Nevertheless, to achieve widespread implementation, future development and implementation efforts must pay special attention to a potential app's ability to meet the individual needs of providers and youths without placing overwhelming burdens on either.

Acknowledgments

We would like to thank our research assistants, Michael Godoy, Jose Padilla-Hernandez, and Dulcinea Jauregui for their hard work and assistance. This research was funded by UCSF Academic Senate, Resource Allocation Program, 7501159.

Conflicts of Interest

None declared.

References

1. Merikangas KR, He J, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010 Oct;49(10):980-989 [FREE Full text] [doi: 10.1016/j.jaac.2010.05.017] [Medline: 20855043]
2. Brauner CB, Stephens CB. Estimating the prevalence of early childhood serious emotional/behavioral disorders: challenges and recommendations. Public Health Rep 2006;121(3):303-310 [FREE Full text] [doi: 10.1177/003335490612100314] [Medline: 16640155]
3. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
4. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. The Lancet 2011 Jun;377(9783):2093-2102. [doi: 10.1016/S0140-6736(11)60512-6]
5. Das JK, Salam RA, Lassi ZS, Khan MN, Mahmood W, Patel V, et al. Interventions for Adolescent Mental Health: An Overview of Systematic Reviews. J Adolesc Health 2016 Oct;59(4S):S49-S60 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.06.020] [Medline: 27664596]
6. Fazel M, Hoagwood K, Stephan S, Ford T. Mental health interventions in schools in high-income countries. The Lancet Psychiatry 2014 Oct;1(5):377-387. [doi: 10.1016/s2215-0366(14)70312-8]
7. García-Carrión R, Villarejo-Carballido B, Villardón-Gallego L. Children and Adolescents Mental Health: A Systematic Review of Interaction-Based Interventions in Schools and Communities. Front Psychol 2019;10:918 [FREE Full text] [doi: 10.3389/fpsyg.2019.00918] [Medline: 31068881]
8. Costello EJ, He J, Sampson NA, Kessler RC, Merikangas KR. Services for adolescents with psychiatric disorders: 12-month data from the National Comorbidity Survey-Adolescent. Psychiatr Serv 2014 Mar 01;65(3):359-366 [FREE Full text] [doi: 10.1176/appi.ps.201100518] [Medline: 24233052]
9. Green JG, McLaughlin KA, Alegría M, Costello EJ, Gruber MJ, Hoagwood K, et al. School mental health resources and adolescent mental health service use. J Am Acad Child Adolesc Psychiatry 2013 May;52(5):501-510 [FREE Full text] [doi: 10.1016/j.jaac.2013.03.002] [Medline: 23622851]
10. Edlund MJ, Wang PS, Berglund PA, Katz SJ, Lin E, Kessler RC. Dropping out of mental health treatment: patterns and predictors among epidemiological survey respondents in the United States and Ontario. Am J Psychiatry 2002 May;159(5):845-851. [doi: 10.1176/appi.ajp.159.5.845] [Medline: 11986140]
11. Cho E, Wood PK, Taylor EK, Hausman EM, Andrews JH, Hawley KM. Evidence-Based Treatment Strategies in Youth Mental Health Services: Results from a National Survey of Providers. Adm Policy Ment Health 2019 Jan;46(1):71-81 [FREE Full text] [doi: 10.1007/s10488-018-0896-4] [Medline: 30209703]
12. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
13. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
14. Shen N, Levitan M, Johnson A, Bender JL, Hamilton-Page M, Jadad AAR, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015 Feb 16;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
15. Anthes E. Mental health: There's an app for that. Nature 2016 Apr 07;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]

16. Abdel-Baki A, Lal S, D-Charron O, Stip E, Kara N. Understanding access and use of technology among youth with first-episode psychosis to inform the development of technology-enabled therapeutic interventions. Early Interv Psychiatry 2017 Feb;11(1):72-76. [doi: 10.1111/eip.12250] [Medline: 26011657]
17. James DCS, Harville C. Smartphone Usage, Social Media Engagement, and Willingness to Participate in mHealth Weight Management Research Among African American Women. Health Educ Behav 2018 Jun;45(3):315-322. [doi: 10.1177/1090198117714020] [Medline: 28606004]
18. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015;6:27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
19. 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] [Medline: 28319239]
20. Aguilera A, Muñoz RF. Text Messaging as an Adjunct to CBT in Low-Income Populations: A Usability and Feasibility Pilot Study. Prof Psychol Res Pr 2011 Dec 01;42(6):472-478 [FREE Full text] [doi: 10.1037/a0025499] [Medline: 25525292]
21. Grist R, Porter J, Stallard P. Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
22. Pramana G, Parmanto B, Kendall PC, Silk JS. The SmartCAT: an m-health platform for ecological momentary intervention in child anxiety treatment. Telemed J E Health 2014 May;20(5):419-427 [FREE Full text] [doi: 10.1089/tmj.2013.0214] [Medline: 24579913]
23. Fleming TM, de Beurs D, Khazaal Y, Gaggioli A, Riva G, Botella C, et al. Maximizing the Impact of e-Therapy and Serious Gaming: Time for a Paradigm Shift. Front Psychiatry 2016;7:65 [FREE Full text] [doi: 10.3389/fpsyt.2016.00065] [Medline: 27148094]
24. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual Research Review: Digital health interventions for children and young people with mental health problems - a systematic and meta-review. J Child Psychol Psychiatry 2017 Apr;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
25. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
26. Massatti RR, Sweeney HA, Panzano PC, Roth D. The de-adoption of innovative mental health practices (IMHP): why organizations choose not to sustain an IMHP. Adm Policy Ment Health 2008 Mar;35(1-2):50-65. [doi: 10.1007/s10488-007-0141-z] [Medline: 17968652]
27. Swain K, Whitley R, McHugo GJ, Drake RE. The sustainability of evidence-based practices in routine mental health agencies. Community Ment Health J 2010 Apr;46(2):119-129. [doi: 10.1007/s10597-009-9202-y] [Medline: 19544094]
28. Castonguay L, Beutler L, editors. Principles of Therapeutic Change That Work. Oxford, UK: Oxford University Press; Sep 15, 2005.
29. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014 Jun 25;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
30. Thomas DR. A General Inductive Approach for Analyzing Qualitative Evaluation Data. American Journal of Evaluation 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
31. Bauer S, Moessner M. Technology-enhanced monitoring in psychotherapy and e-mental health. J Ment Health 2012 Aug;21(4):355-363. [doi: 10.3109/09638237.2012.667886] [Medline: 22548363]
32. Lambert MJ, Whipple JL, Hawkins EJ, Vermeersch DA, Nielsen SL, Smart DW. Is It Time for Clinicians to Routinely Track Patient Outcome? A Meta-Analysis. Clinical Psychology: Science and Practice 2003;10(3):288-301. [doi: 10.1093/clipsy.bpg025]
33. Carstens CA, Panzano PC, Massatti R, Roth D, Sweeney HA. A naturalistic study of MST dissemination in 13 Ohio communities. J Behav Health Serv Res 2009 Jul;36(3):344-360. [doi: 10.1007/s11414-008-9124-4] [Medline: 18663581]
34. 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] [Medline: 30600619]
35. Mackenzie CS, Gekoski WL, Knox VJ. Age, gender, and the underutilization of mental health services: the influence of help-seeking attitudes. Aging Ment Health 2006 Nov;10(6):574-582. [doi: 10.1080/13607860600641200] [Medline: 17050086]

Abbreviations

mHealth: mobile health

Edited by J Torous; submitted 21.09.20; peer-reviewed by J Manning, D Szinay; comments to author 03.12.20; revised version received 26.01.21; accepted 22.02.21; published 22.04.21.

Please cite as:
Armstrong CC, Odukoya EJ, Sundaramurthy K, Darrow SM
Youth and Provider Perspectives on Behavior-Tracking Mobile Apps: Qualitative Analysis
JMIR Ment Health 2021;8(4):e24482
URL: https://mental.jmir.org/2021/4/e24482
doi:10.2196/24482
PMID:33885364

Review

Digital Health Interventions in Prevention, Relapse, and Therapy of Mild and Moderate Depression: Scoping Review

Pinar Tokgöz[1*], MPH; Robert Hrynyschyn[2*], MPH; Jessica Hafner[1*], MPH; Simone Schönfeld[1,3,4], MPH; Christoph Dockweiler[1*], Prof Dr

[1]School of Public Health, Bielefeld University, Bielefeld, Germany

[2]Charité – Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Health and Nursing Science, Berlin, Germany

[3]LWL-Klinik Lippstadt und Warstein, Lippstadt, Germany

[4]Universität Witten/Herdecke, Institut für Integrative Gesundheitsversorgung und Gesundheitsförderung, Witten, Germany

[*]these authors contributed equally

Corresponding Author:
Pinar Tokgöz, MPH
School of Public Health
Bielefeld University
Universitätsstraße 25
Bielefeld, 33615
Germany
Phone: 49 521 106 67136
Email: pinar.tokgoez@uni-bielefeld.de

Abstract

Background: Depression is a major cause for disability worldwide, and digital health interventions are expected to be an augmentative and effective treatment. According to the fast-growing field of information and communication technologies and its dissemination, there is a need for mapping the technological landscape and its benefits for users.

Objective: The purpose of this scoping review was to give an overview of the digital health interventions used for depression. The main goal of this review was to provide a comprehensive review of the system landscape and its technological state and functions, as well as its evidence and benefits for users.

Methods: A scoping review was conducted to provide a comprehensive overview of the field of digital health interventions for the treatment of depression. PubMed, PSYNDEX, and the Cochrane Library were searched by two independent researchers in October 2020 to identify relevant publications of the last 10 years, which were examined using the inclusion and exclusion criteria. To conduct the review, we used Rayyan, a freely available web tool.

Results: In total, 65 studies were included in the qualitative synthesis. After categorizing the studies into the areas of prevention, early detection, therapy, and relapse prevention, we found dominant numbers of studies in the area of therapy (n=52). There was only one study for prevention, 5 studies for early detection, and 7 studies for relapse prevention. The most dominant therapy approaches were cognitive behavioral therapy, acceptance and commitment therapy, and problem-solving therapy. Most of the studies revealed significant effects of digital health interventions when cognitive behavioral therapy was applied. Cognitive behavioral therapy as the most dominant form was often provided through web-based systems. Combined interventions consisting of web-based and smartphone-based approaches are increasingly found.

Conclusions: Digital health interventions for treating depression are quite comprehensive. There are different interventions focusing on different fields of care. While most interventions can be beneficial to achieve a better depression treatment, it can be difficult to determine which approaches are suitable. Cognitive behavioral therapy through digital health interventions has shown good effects in the treatment of depression, but treatment for depression still stays very individualistic.

(JMIR Ment Health 2021;8(4):e26268) doi:10.2196/26268

KEYWORDS

digital health; depression; scoping review; health care

Introduction

Depressive disorders are among the most significant and widespread diseases, and their relevance will continue to increase in the coming years. Nearly 300 million people are affected worldwide [1]. Depressive disorders have a high risk of chronicity and are associated with a substantially increased probability of developing further comorbidities with corresponding effects on the quality of life of those affected. Psychiatric care is facing different structural problems of care, with underuse and misuse of care that is evident in practice [1]. Barriers to access to psychiatric specialists or psychotherapeutic care, lack of cross-sectoral and interdisciplinary care, and long waiting lists are also challenging. It is estimated that less than 50% of those currently affected receive therapy that is appropriate to their needs and requirements according to current scientific criteria. Although there is an increasing awareness of mental health issues, accessibility to health care has been a key problem [2].

Against this background, psychological, psychotherapeutic, and psychiatric care are experiencing remarkable developments in technology-supported care concepts. A broad spectrum of digital health interventions (DHIs) in psychiatric care already exist today [3]. DHIs enable new forms of interaction and knowledge-based reproduction in the field of health. The constantly growing number of interventions extends from outpatient digital health care programs to telephone, video, or text-based interactions with the therapist and complex online-based intervention programs. Unlike face-to-face treatment, such support systems are easily accessible and standardized, and they can reduce the fear of stigmatization, as they can be used in private and at the convenience of the patient [3].

A steadily increasing amount of empirical data show first indications of patient-related benefits of DHIs, especially the reduction of depressive symptoms; improvement in quality of life; and reduction of direct, indirect, and intangible costs [3]. These potential benefits could be effective for patients and the health system if successful acceptance of DHIs is achieved [4].

This scoping review is part of a research project that examines the multiperspective and participatory development of technology-supported care for people with depressive disorders. The purpose of this scoping review was to give an overview of DHIs used in different fields of depression care. The main goal was to provide a comprehensive review of the system landscape and its technological state and functions, as well as the evidence and benefits for users.

To this end, the following research questions have been addressed:

- What types of DHIs for the treatment of mild and moderate depression have been developed, and how can the functions be described?
- How can the benefits of DHIs in the care of mild and moderate depression be described?

Methods

Overview

A scoping review has been conducted to identify knowledge gaps, set research agendas, and identify implications for intervention development. Although scoping reviews are related to systematic reviews, they differ in numerous ways. Scoping reviews present a broader overview of evidence pertaining to a specific topic, irrespective of study quality [5], useful when emerging topics are discussed to clarify key concepts and research gaps. Systematic reviews focus more on specific research questions with a priori defined criteria. Therefore, scoping reviews generate hypotheses, while systematic reviews focus more on testing hypotheses [6]. Results were reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines of systematic reviews because there are no reporting guidelines for scoping reviews. Research questions and inclusion criteria were adapted from Arksey and O'Malley [5].

Inclusion and Exclusion Criteria

The following criteria have been developed in coordination with the partners of the research project. The inclusion criteria comprised (1) all DHIs for the treatment of depression, (2) quantitative study design, (3) language: English or German, and (4) published between 2010 and 2020 since the technological development of DHI is dynamic. Exclusion criteria were (1) study participants aged younger than 18 years, (2) no diagnosis of depression, (3) severe course of illness, (4) psychotic symptoms, (5) concurrent medication therapy, or (6) DHIs with no clear relation to depression treatment.

Literature Search

The search was conducted in October 2020 in the databases PubMed, PSYNDEX, and the Cochrane Library. The following search strategy was used: (depression OR depressive disorder OR depressive episode) AND (online based* OR mobile* OR ehealth* OR "electronic mental health" OR "e-mental health" OR online-based* OR internet-based* OR web-based* OR computer-based).

Data Extraction

The articles were extracted using standardized table formats. To provide an overview of the aspects of DHIs for depression considered here, the following taxonomy is presented in the tables:

- Authors, year, country, and funding of the study
- Study design
- Study period
- Sample size information
- Technology description and functions
- Relevant outcomes and effects

Moreover, the results were analyzed regarding the effects of the reduction of depressive symptoms and specific characteristics of the DHI. For this purpose, contingency tables were used to show frequency distributions between the benefit and the technology used and the benefit and the therapy form

used. Furthermore, graphs were built to show which functions have been used by the different therapy forms.

Results

Study Selection

In total, 3040 publications were identified. These publications were transmitted to Rayyan, which is a free web tool designed to help researchers conducting systematic reviews. Using this program, the 2 independent researchers (PT and RH) screened the articles according to inclusion and exclusion criteria. In case of disagreement, consensus has been made by the opinion of a third researcher (JH). A total of 65 studies were included in the qualitative synthesis (Figure 1).

In a first step, the field of depression care was categorized in 4 application areas: prevention (1 study identified), early detection (5 studies), therapy (52 studies), and relapse prevention (7 studies; Figure 2). In a further step, this standardization enabled technology mapping and a detailed description of the technologies and their benefits.

Figure 1. PRISMA flow diagram of literature search and selection process [7].

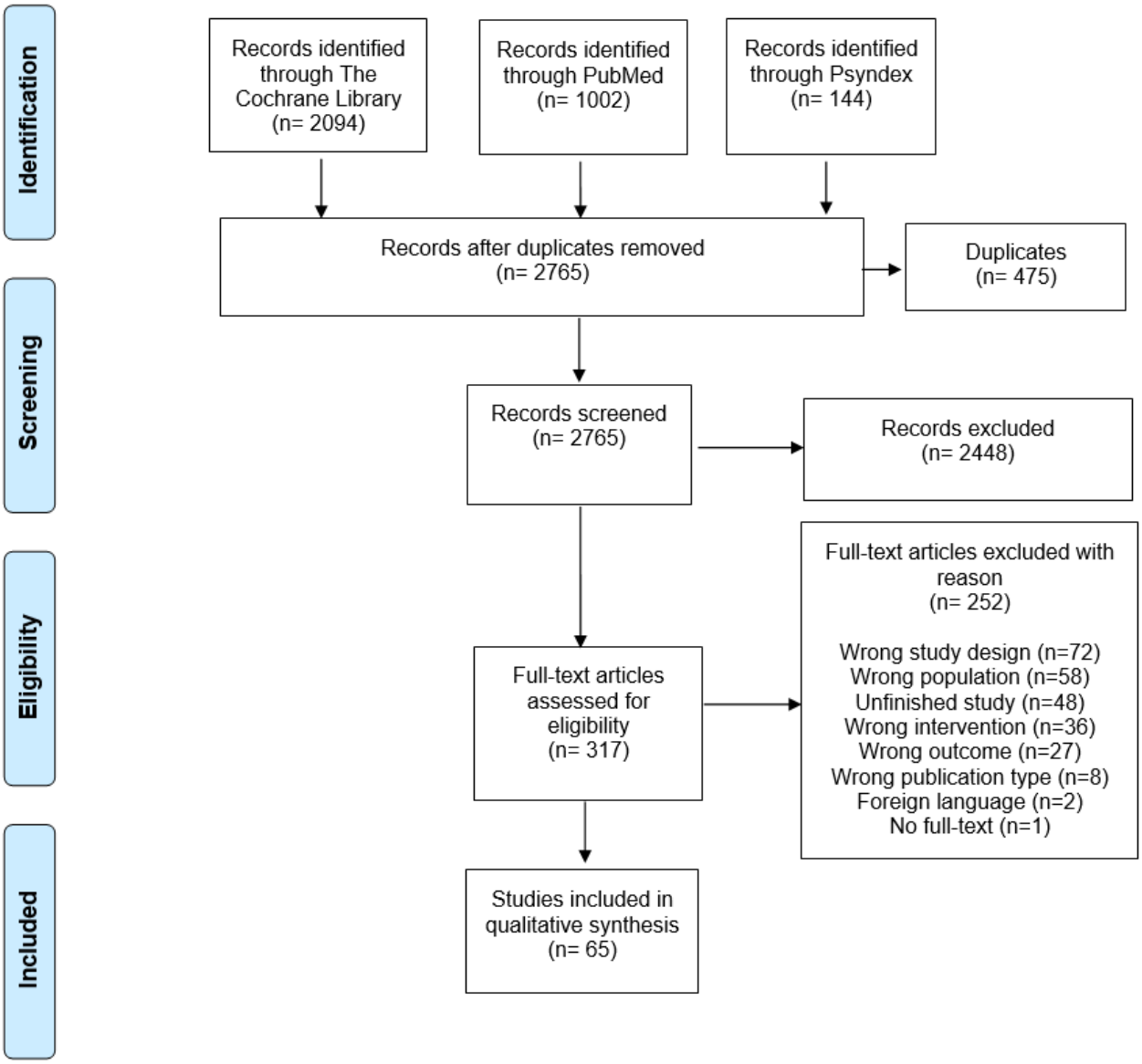
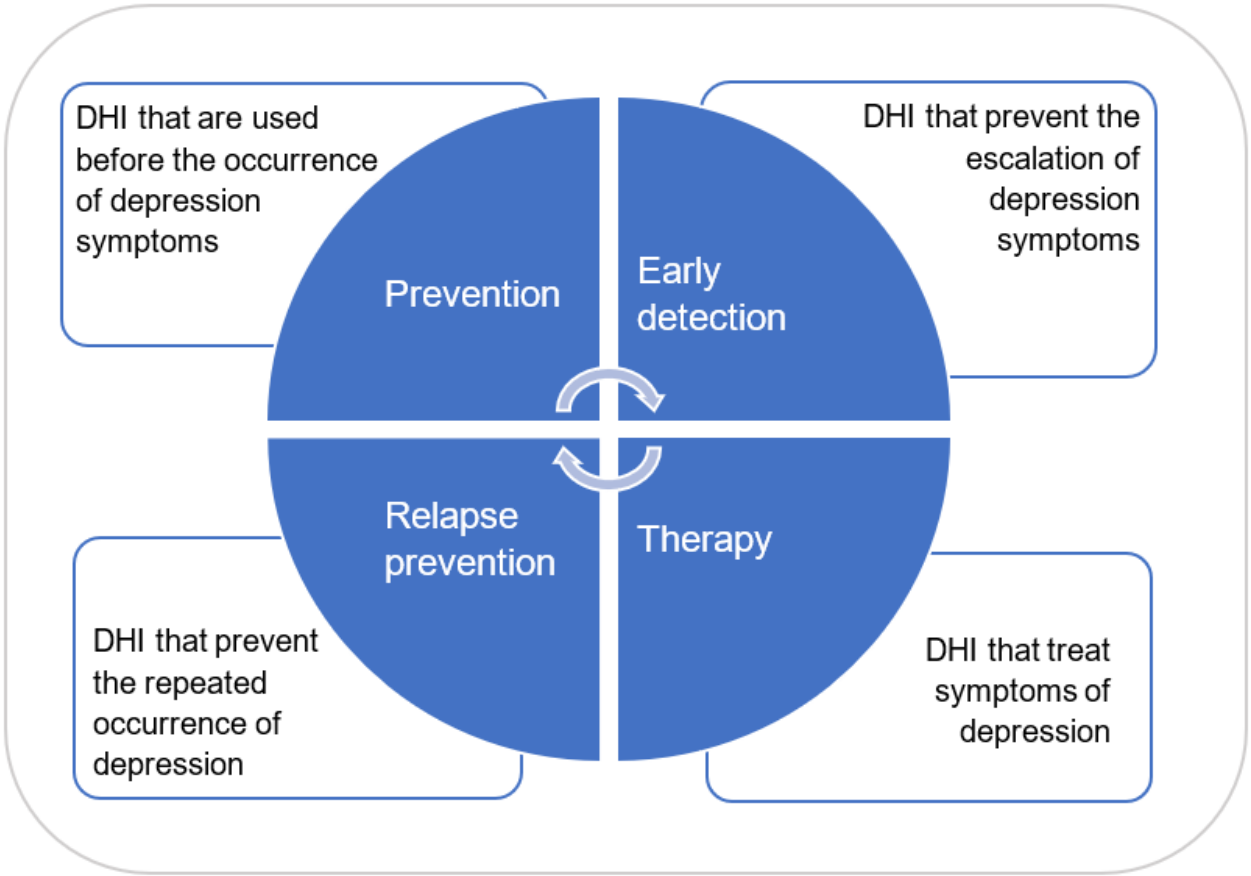


Figure 2. Classification of digital health intervention application areas.



Data Description: Prevention/Early Detection

In the area of prevention, only one study was available in context of this scoping review [8]. The study was conducted in Germany and funded by the European Union (EU). Its duration was 6 weeks, with follow-ups after 6 months and 12 months. The study population consists of people with underlying depressive symptoms. Differentiation between an intervention group (IG) and control group (CG) was made (IG=202 and CG=204; Multimedia Appendix 1).

The intervention GET.ON Mood Enhancer Prevention is a web-based guided self-help intervention based on psychoeducation, problem-solving therapy, and behavioral activation. The concepts were conveyed via multimodal and interactive elements in 6 sessions of about 30 minutes each with individual feedback by an online trainer. The CG received psychoeducation via the same web-based platform, but without the guidance of an online trainer. The primary outcome of the study was the diagnosis of major depression, which was recorded using the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5) criteria and a Structured Clinical Interview for DSM-5 (SCID).

Five studies were identified in the area of early detection (Multimedia Appendix 1). Two of the studies were conducted in Germany [9,10], 2 in the United States [11,12], and 1 in Romania [13]. Two studies were financed by the EU [9,10], and 3 were funded by countries [11-13].

In all 5 studies, the study populations consisted of people with underlying depressive symptoms (ie, no manifest depressive disorder). In 4 studies, a differentiation was made between the IG and CG. The number of participants ranged from 35 to 202 in the IG (mean 102; mode 85.5) and from 36 to 204 participants in the CG (mean 93.7; mode 69). One study was a comparison of different interventions and did not have a CG [12]. The study periods ranged from 3 to 24 weeks.

Three out of the 5 studies chose a mixed form as treatment approach. In 2 studies, the intervention consisted of web-based cognitive behavioral therapy (CBT) and problem-solving therapy [9,10]. One study consisted of an app-based intervention based on CBT and problem-solving therapy, as well as an information control app [12]. In all 3 studies, the participants of the IG received support or reminder from an online trainer via a news system. Optionally, the participants could also choose to receive motivational text messages regularly to ensure continuity. In Place et al [11], the intervention consisted of a monitoring system that collected different types of metadata and clinicians providing feedback. In Buntrock et al [9], the CG only received psychoeducation through the web-based platform. In 2 studies, participants of the CG had unlimited access to treatment as usual (TAU) during the study period and were given access to the web-based intervention after the study period [10,11].

The study by Tulbure et al [13] investigated a web-based transdiagnostic intervention. Thereby, therapy concepts of different mental diseases were transferred. In this study,

concepts from the field of anxiety disorders were projected onto the field of depressive disorders. The IG received a web-based multimodal intervention in which additional writing tasks were completed. The participants could perform the intervention on a computer, tablet, or smartphone. Trained personnel monitored the activity of the participants. In addition, there was personalized feedback on writing tasks. Furthermore, participants received a reminder if inactivity was detected. No information was provided regarding the therapeutic approach. The CG received reading access to the platform [13].

All studies described depression-specific symptoms as primary outcome [9-13]. Additionally, in 2 studies, adherence to the intervention [9,13] was indicated; in 1 study, quality of life [13] was.

Data Description: Relapse Prevention

In the area of relapse prevention, 4 studies from Germany [14-17], 1 study from the Netherlands [18], 1 study from Denmark [19], and 1 from the United States [20] were identified (Multimedia Appendix 1).

The study periods ranged from 4 weeks to 104 weeks (24 months). Four studies were financed by countries [14,17,18,20], and 2 studies were privately financed [16,19]. One study did not provide information on financing [15]. Six studies differentiated between an IG and a CG [14-18,20]. The studies revealed large differences in sample size, ranging from 41 to 264 participants (mean 182; mode 217.5) when the IG and CG were combined. The sample sizes in the experimental arms of the studies varied between 21 and 230 participants (mean 115.2; mode 120). In the CG, the samples varied between 20 and 230 participants (mean 101.2; mode 91.5).

The interventions showed differences in their implementation. While almost all of them used a feedback function to indicate therapy success or remind patients to perform tasks, the therapy strategy differed. Four out of the 7 studies chose CBT as a therapeutic approach [16-18,20]. Some studies used CBT, either through a telerehabilitation program [16], mobile app [18], or web-based program [20]. Kraft et al [14] opted for a mindfulness-based exercise in which patients received feedback on the success via SMS. Lauritsen et al [19] used a web-based self-assessment tool that monitored participant mood, sleep, and activity, and Zwerenz et al [17] used a web-based self-help program based on CBT. One study did not provide any information on the chosen therapeutic approach [15]. Instead of the technical intervention, the CG of all studies received the nontechnical comparative therapy (ie, a comparable therapy) performed without the use of technical aid or TAU.

Five studies described depression-specific symptoms as primary outcome [15-17,19,20]. Additionally, adherence to the intervention [14,19] and quality of life [17,19] were indicated.

Data Description: Therapy

For the area of therapy, 52 studies were identified. A total of 12 studies were conducted in the United States [21-32], 6 studies in the Netherlands [33-38], and 11 in Germany [39-49]. Four studies were conducted in Australia [50-53]. Three studies were conducted in Switzerland [54-56] and 4 in Spain [57-60]. In Canada [61,62], Finland [63,64], Sweden [65,66], and Great Britain [67,68] 2 studies each were identified. New Zealand [69], Austria [70], Ireland [71], and Japan [72] were each represented with 1 study (Multimedia Appendix 1).

Of those, 27 were sponsored by countries [23,25,26,28-32,35,40,41,43,44,46,51,52,57-62,65,66,68,69,71] and 11 studies were privately financed [24,33,34,36,37,48,49,53,55,56,67]. The EU funded 2 studies [39,42]. One study indicated a mixture of private and state funding [21]. Eleven studies did not provide any information of financial sources [22,27,38,45,47,50,54,63,64,70,72].

In the studies differentiating between an IG and CG, the number of participants ranged from 10 to 1904 participants in the IG (mean 191.0; mode 88) and from 8 to 1901 participants in the CG (mean 179.1; mode 67.5). The intervention periods of the studies ranged from 2 to 52 weeks (mean 11 weeks).

The majority of studies (41/52) differentiated between an IG and CG and were conducted as randomized controlled trials (RCTs) [22,24,27-29,33-52,54,55,57-61,63-69,71,72]. Seven studies were conducted as pilot studies [25,26,30-32,56,62]. Two studies had a quasi-experimental design [23,70]. One study was conducted as a usability study [53] and 1 as a controlled clinical trial [21].

Technology Mapping: Therapy

The therapy area yielded the largest number of studies. Therefore, the following technology mapping and analysis focus on therapeutic interventions for the reduction of depressive symptoms. In terms of therapeutic approach, CBT, acceptance and commitment therapy (ACT), and problem-solving therapy was considered because those are the most used approaches, and there was sufficient evidence to analyze. Studies were examined according to the type of technology used, medium used, and functions offered. Cross-references between the levels were drawn with the help of cross tables.

CBT was the dominant form of therapy in 54% (28/52) of the studies [21,23,24,28,30,32,39-41,43-46,48-50,52,55-57,60-62,65-67,70-72]. In 19% (9/52), mixed forms were used, consisting of 2 or more forms of therapy [27,31,33,34,42,47,51,58,59]. Other therapy approaches such as ACT (5/52, 9%) [26,35,38,63,64] and problem-solving therapy (4/52, 8%) [29,36,37,69] were significantly less common in interventions. The remaining studies either focused on cognitive restructuring [54], cognitive control therapy [22], behavioral activation [45,68], or supplied no information about the therapeutic approach that was used [25].

Considering how therapy was technologically implemented, differences become apparent. Most of the studies (39/52, 75%) used a web-based system in which an application was used online [21-23,28,33-46,48,49,52-55,57-67,69-72]. This form of therapy provision was the most common form. Mixed forms were also frequently used (5/52, 11%) using web-based and smartphone-based approaches [24,27,30,32,68]. Smartphone-based approaches were used by 12% (6/52) of studies [25,26,31,47,50,56]. One study (2%) featuring problem-solving therapy used virtual reality for delivery, which

entailed a virtual therapist giving instructions on skills in problem solving [29].

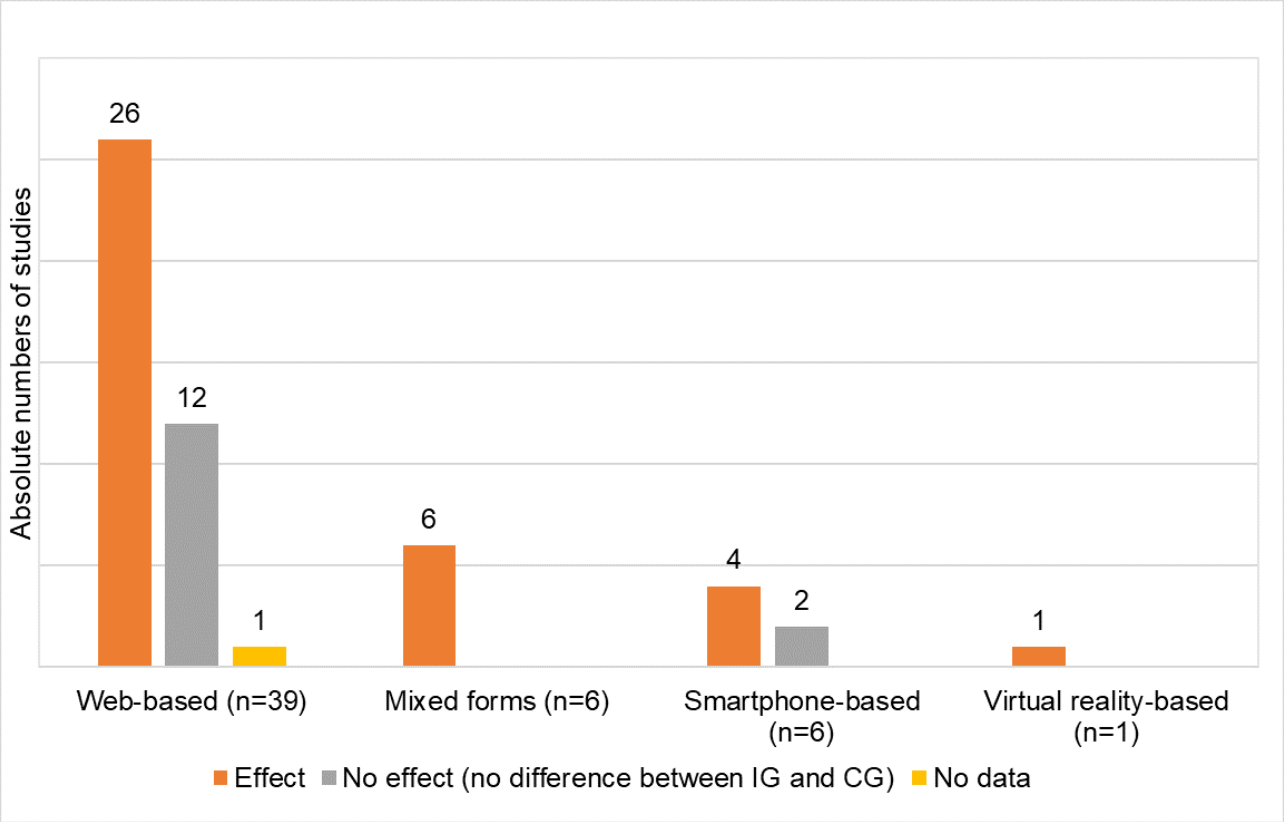
In addition, we analyzed which functions were integrated in the DHI. Multiple answers were possible, depending on which functions were given. Most of the DHIs (33/52, 63%) were offered as guided interventions in which a therapist supported the participants [21,25,26,28,29,31-39,42-44,47,49,54,55,57-60,62-64,67-71], with 15% (11/52) conducted as unguided interventions [22,23,27,30,40,45,46,49,50,52,61] and 15% (11/52) using the reminder function to encourage patients to perform daily exercises or self-assessments [24,32,41,47,49,51,63-65,67,68]. While some studies (6/52, 8%) used tracking functions with biosensors [24,50,51,53,56,59], 3% (2/52) used diary functions [50,51] and 3% (2/52) used discussion forums [32,62]. A total of 9% (5/52) of studies used educational elements for the treatment of depression [53,56,66,70,72].

Evidence of DHI in Therapy for Depression

A total of 71% (37/52) of studies reported improvements in depressive symptoms [22,30,32,37,40,42,44,46,49-51,53-57,59-64,67,68,70-72], 27% (14/52) reported no (significant) effects [21,31,33-36,41,45,47,48,58,65,66,69], and 2% (1/52) did not consider depressive symptoms as an outcome [52].

In Figure 3, the type of technology used with reported benefits of the studies is shown. In the case of web-based provision of DHIs, improvements of depressive symptoms becomes clear. A total of 67% (26/39) of studies identified a benefit [22,23,28,37,40,42,44,46,49,52,55,57,59-64,67,70-72], 31% (12/39) did not verify an improvement in depressive symptoms [21,33-36,41,45,48,58,65,66,69], and 1 study did not provide results for depressive outcomes [52]. All studies that used mixed forms of technology demonstrated a benefit. However, due to the small number of studies, this should be interpreted cautiously.

Figure 3. Evidence for digital health interventions classified by technology.



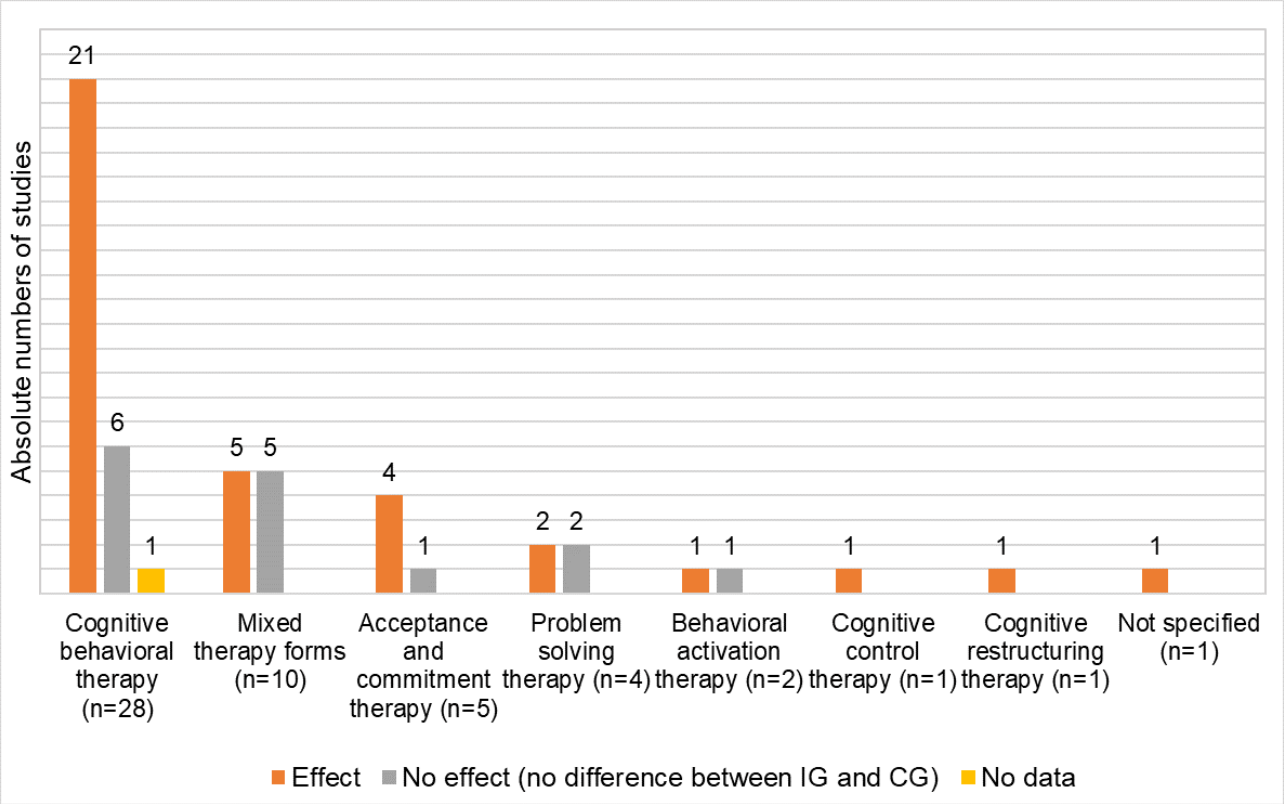
Evidence for CBT

In Figure 4, the therapy approach used with reported benefits is shown. A total of 75% (21/28) of studies that chose a CBT approach reported a positive effect on the reduction of depressive symptoms [23,24,28,30,32,39,40,44,46,49,50,55,57,60,62,67,70-72], 21% (6/28) could not find any effect in terms of reduction [21,41,43,47,65,66], and 4% (1/28) did not report any information on the depression outcome [52].

Of the studies with a positive effect, 76% (16/21) were transmitted via web-based applications [23,28,39,40,44,46,49,55,57,60-62,67,70-72], 48% (10/21) were designed as guided [28,39,44,55,57,60,62,67,70,71] and 29% (6/21) as unguided interventions [23,40,46,49,61,72].

In contrast, 6 studies could be identified that showed no or neutral results: 5 were web-based interventions [21,41,43,65,66] and 1 study was designed as an app [47].

Figure 4. Evidence of digital health interventions classified by therapeutic approach.



Evidence for Problem-Solving Therapy

The evidence for benefits using the problem-solving therapy approach is controversial. Of the studies identified, 50% (2/4) determined a benefit in terms of an improvement in depressive symptoms [29,37] and 50% (2/4) could not determine a difference between the IG and CG or simply do not represent significant improvement in the IG [36,69]. One of the studies with a reported benefit was a web-based application [37] and the second was a virtual reality intervention [29].

Evidence for ACT

Of the studies that chose ACT as a therapeutic approach, 80% (4/5) demonstrated a benefit regarding depressive symptoms [26,38,63,64]. One of the studies with a reported benefit was an app-based intervention [26] and 3 were designed as guided web-based interventions [38,63]. The study that did not indicate a benefit was a web-based intervention with automated feedback function [35].

Discussion

Principal Findings

The number of DHIs for treating depression is quite comprehensive. Focusing solely on therapeutic setting is not sufficient to show the different approaches of depression care. However, most of the studies were found in the therapy area. Of the 65 studies included, 52 focused on a therapy setting, 7 on relapse prevention, 5 on early detection, and 1 on the prevention of depression.

The most dominant approaches were CBT, ACT, and problem-solving therapy. Regarding the efficacy, most of the studies finding a benefit regarding symptom severity used CBT. Most studies found a significant effect of digital CBT for depression symptoms. Other studies not using DHIs for the provision of CBT have similar results for the efficacy of CBT regarding reduction of depression symptoms. Especially for mild and moderate depression, there is good evidence for symptom reduction [73]. Beyond that, the DHIs in most of the identified studies were realized through web-based technologies; mobile or smartphone-based interventions were still underrepresented.

Most of the included studies compared participants treated with DHIs with those not treated at all (eg, waitlist). In contrast, there were only a few studies that compared the efficacy of DHIs with traditional psychotherapy. It is not surprising that DHIs work better than no treatment at all. However, their efficacy is quite remarkable. Additionally, most studies were conducted under ideal controlled conditions, and studies investigating effectiveness in everyday life are still missing. For this reason, there are emerging discussions about adaptive study designs such as n-of-1 or interrupted time-series designs, which can be better suited to evaluate DHIs. Although decision makers or payers still prefer RCTs as the gold standard for evidence, there are limitations when it comes to the measurement of efficacy and effectiveness for DHIs. Although health insurance companies offer their insureds DHIs, only 3% to 25% of patients take advantage of them [74]. Reasons for this are low expectations of their effectiveness, concerns about data security, poor user-friendliness, general skepticism about psychotherapy, and little experience with the internet in general [74].

Comparison With Prior Work

However, even for CBT, which is increasingly delivered by computerized forms, there is good evidence supporting the stated effects in this study. In the meta-analysis of Andersson and Cuijpers [75] that examined the effects of computerized CBT compared with face-to-face interventions, the authors found that computerized treatments do have promising potential for treating depression, and the computerized treatments were statistically significantly superior (posttest effect size: 0.41, 95% CI 0.29-0.54, I^2=57%) to face-to-face interventions.

The dissemination of technological systems poses a challenge for the field. DHIs developed within a research context vary in their implementation and technological approaches. Differences in technological attributes have been discussed in the literature. Studies have consistently demonstrated that effects appear to increase with higher levels of human guidance (eg, no guidance vs administrative guidance vs therapist guidance) [4,76]. Johansson et al [77] showed in a systematic review of 25 controlled trials that guided internet-based CBT is more effective than unguided internet-based CBT. Whereas the effect sizes in studies with people who had no therapist contact either before or during the treatment were relatively low (average effect size d=0.21), they improved with more support from the therapist. Studies with people who had contact with a therapist before and during the treatment reported a much higher average effect size (d=0.76). Wright et al [78] showed similar results in a recent study from 2019. In a systematic review with 40 RCTs, they showed that studies providing support by a therapist or clinician yielded larger effects (g=0.67) than studies with unguided interventions (g=0.24). One reason for this could be the missing therapeutic alliance was not compensated for with more content or technological support, apart from reminders. Therefore, it is possible that guidance becomes important, especially when systems are not very responsive.

Aside from this, there are many opportunities for DHIs in the treatment of depression. In fact, interventions are mostly delivered on the internet. Only a few publications considering mobile interventions were identified. This is surprising since smartphones are ubiquitous and highly prevalent nowadays. A review from 2015 identified 82 mobile apps for depression treatment [79]. Another review found that only 5 of those apps had been empirically evaluated in RCTs [80]. Regarding the high diffusion of commercially driven apps, scientifically evaluated apps are needed in future. In addition, innovative technologies using social media or virtual reality are underrepresented in this review.

Limitations

Most of the studies were found in the area of therapy. One reason for this could be that preventive offers are not recognized the way therapeutic interventions are, even if those preventive interventions are advisable due to individual risk factors. It could be helpful to provide a wider range of preventive services and a larger number of psychotherapists as well as general practitioners to give their patients recommendations for preventive interventions.

All studies tried to determine the effects of depression-specific symptoms, but in doing so the authors of the studies used different instruments. While some researchers used the Center for Epidemiological Studies Depression Scale or the Quick Inventory of Depressive Symptomatology–16 item scale, the Patient Health Questionnaire–9 Item and Beck Depression Inventory II scales were most commonly used. Because of this heterogeneity, there is a limitation considering the comparability of studies.

A limitation concerns the scope of the reviewed publications. Systems developed for children and adolescents and women with postpartum depression and studies published before 2010 were excluded. Therefore, it is necessary to investigate in further research if the findings can be generalized to those types of systems. As data security is an important concern of DHIs, consistent guidelines and criteria for future systems are needed. Despite these limitations, the findings of our analysis highlight some of the challenges and opportunities for the use of DHIs for depression.

Conclusion

The aim of this scoping review was to approach the state of DHIs for mild and moderate depression and try to describe those interventions for depression care. It is known that DHIs have great potential to complement the treatment of depressive people and intensify traditional psychotherapy. This review indicates DHIs also show great potential in avoiding depression (prevention and early detection) and preventing relapse, which is currently underrepresented in the literature. The field of DHIs is constantly increasing, and there is no general evidence of the effect of these interventions, but most of the studies included in this review improved the effects of DHIs for the reduction of depressive symptoms. Furthermore, results showed that most of the DHIs provided web-based CBT and a combination of several technological options (eg, reminder, tracking). Even with automated programs and contact with psychotherapists (eg, contact through chats or email), people show great improvement in their symptoms in comparison to interventions without support at all. It appears to be crucial for people to have professional guidance. Further research is needed to investigate whether more technologically advanced systems lead to higher adherence and effects for the reduction of depressive symptoms in addition to investigating real-world effectiveness and barriers for implementation of DHIs.

Acknowledgments

We acknowledge support for the publication costs from the Deutsche Forschungsgemeinschaft and the Open Access Publication Fund of Bielefeld University. This project was financially supported by a grant from the Landeszentrum Gesundheit North Rhine-Westphalia (LZG.NRW), a professional and state-run coordination center for public health services in North Rhine-Westphalia within the portfolio of the Ministry of Labour, Health, and Social Affairs of North Rhine-Westphalia that gives

support and expert advice in all health issues to the authorities in the government and municipalities. The LZG.NRW played no role in the planning and conceptualization of the manuscript, did not participate in the writing of the manuscript, and played no role in the decision to submit the manuscript.

Authors' Contributions

PT, RH, JH, SS, and CD were involved and contributed at each step of manuscript writing. All the authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study characteristics of the included studies.
[DOCX File , 58 KB - mental_v8i4e26268_app1.docx]

References

1. Herrman H, Kieling C, McGorry P, Horton R, Sargent J, Patel V. Reducing the global burden of depression: a Lancet-World Psychiatric Association Commission. Lancet 2019 Jun 15;393(10189):e42-e43. [doi: 10.1016/S0140-6736(18)32408-5] [Medline: 30482607]
2. Rycroft-Malone J, Gradinger F, Owen Griffiths H, Anderson R, Crane RS, Gibson A, et al. "Mind the gaps": the accessibility and implementation of an effective depression relapse prevention programme in UK NHS services: learning from mindfulness-based cognitive therapy through a mixed-methods study. BMJ Open 2019 Sep 08;9(9):e026244 [FREE Full text] [doi: 10.1136/bmjopen-2018-026244] [Medline: 31501097]
3. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
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] [Medline: 22466510]
5. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
6. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping review on the conduct and reporting of scoping reviews. BMC Med Res Methodol 2016 Feb 09;16:15 [FREE Full text] [doi: 10.1186/s12874-016-0116-4] [Medline: 26857112]
7. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
8. Buntrock C, Ebert DD, Lehr D, Smit F, Riper H, Berking M, et al. Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: a randomized clinical trial. JAMA 2016 May 3;315(17):1854-1863. [doi: 10.1001/jama.2016.4326] [Medline: 27139058]
9. Buntrock C, Ebert D, Lehr D, Riper H, Smit F, Cuijpers P, et al. Effectiveness of a web-based cognitive behavioural intervention for subthreshold depression: pragmatic randomised controlled trial. Psychother Psychosom 2015;84(6):348-358. [doi: 10.1159/000438673] [Medline: 26398885]
10. Ebert DD, Buntrock C, Lehr D, Smit F, Riper H, Baumeister H, et al. Effectiveness of web- and mobile-based treatment of subthreshold depression with adherence-focused guidance: a single-blind randomized controlled trial. Behav Ther 2018 Dec;49(1):71-83. [doi: 10.1016/j.beth.2017.05.004] [Medline: 29405923]
11. Place S, Blanch-Hartigan D, Smith V, Erb J, Marci CD, Ahern DK. Effect of a mobile monitoring system vs usual care on depression symptoms and psychological health: a randomized clinical trial. JAMA Netw Open 2020 Jan 03;3(1):e1919403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.19403] [Medline: 31951270]
12. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. J Med Internet Res 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]
13. Tulbure BT, Rusu A, Sava FA, Sălăgean N, Farchione TJ. A web-based transdiagnostic intervention for affective and mood disorders: randomized controlled trial. JMIR Ment Health 2018 May 24;5(2):e36 [FREE Full text] [doi: 10.2196/mental.8901] [Medline: 29798831]
14. Kraft S, Wolf M, Klein T, Becker T, Bauer S, Puschner B. Text message feedback to support mindfulness practice in people with depressive symptoms: a pilot randomized controlled trial. JMIR Mhealth Uhealth 2017 May 02;5(5):e59 [FREE Full text] [doi: 10.2196/mhealth.7095] [Medline: 28465278]

15. Schlicker S, Ebert DD, Middendorf T, Titzler I, Berking M. Evaluation of a text-message-based maintenance intervention for Major Depressive Disorder after inpatient cognitive behavioral therapy. J Affect Disord 2018 Feb;227:305-312. [doi: 10.1016/j.jad.2017.10.047] [Medline: 29132073]
16. Schmadeke S, Bischoff C. Effects of smartphone-supported rehabilitation aftercare (eATROS) for depressive patients. Verhaltenstherapie 2015 Nov 23;25(4):277-286. [doi: 10.1159/000441856]
17. Zwerenz R, Baumgarten C, Becker J, Tibubos A, Siepmann M, Knickenberg RJ, et al. Improving the course of depressive symptoms after inpatient psychotherapy using adjunct web-based self-help: follow-up results of a randomized controlled trial. J Med Internet Res 2019 Oct 24;21(10):e13655 [FREE Full text] [doi: 10.2196/13655] [Medline: 31651403]
18. Klein NS, Bockting CL, Wijnen B, Kok GD, van Valen E, Riper H, et al. Economic evaluation of an internet-based preventive cognitive therapy with minimal therapist support for recurrent depression: randomized controlled trial. J Med Internet Res 2018 Dec 26;20(11):e10437 [FREE Full text] [doi: 10.2196/10437] [Medline: 30478021]
19. Lauritsen L, Andersen L, Olsson E, Søndergaard SR, Nørregaard LB, Løventoft PK, et al. Usability, acceptability, and adherence to an electronic self-monitoring system in patients with major depression discharged from inpatient wards. J Med Internet Res 2017 Dec 21;19(4):e123 [FREE Full text] [doi: 10.2196/jmir.6673] [Medline: 28432040]
20. Segal ZV, Dimidjian S, Beck A, Boggs JM, Vanderkruik R, Metcalf CA, et al. Outcomes of online mindfulness-based cognitive therapy for patients with residual depressive symptoms: a randomized clinical trial. JAMA Psychiatry 2020 Jun 01;77(6):563-573. [doi: 10.1001/jamapsychiatry.2019.4693] [Medline: 31995132]
21. 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] [Medline: 28483742]
22. Calkins AW, McMorran KE, Siegle GJ, Otto MW. The effects of computerized cognitive control training on community adults with depressed mood. Behav Cogn Psychother 2015 Sep;43(5):578-589. [doi: 10.1017/S1352465814000046] [Medline: 24589123]
23. Dimidjian S, Beck A, Felder JN, Boggs JM, Gallop R, Segal ZV. Web-based mindfulness-based cognitive therapy for reducing residual depressive symptoms: an open trial and quasi-experimental comparison to propensity score matched controls. Behav Res Ther 2014 Dec;63:83-89. [doi: 10.1016/j.brat.2014.09.004] [Medline: 25461782]
24. Hirsch A, Luellen J, Holder JM, Steinberg G, Dubiel T, Blazejowskyj A, et al. Managing depressive symptoms in the workplace using a web-based self-care tool: a pilot randomized controlled trial. JMIR Res Protoc 2017 Apr 04;6(4):e51 [FREE Full text] [doi: 10.2196/resprot.7203] [Medline: 28377368]
25. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. J Med Internet Res 2017 Jan 05;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
26. Pfeiffer PN, Henry J, Ganoczy D, Piette JD. Pilot study of psychotherapeutic text messaging for depression. J Telemed Telecare 2017 Aug;23(7):665-672. [doi: 10.1177/1357633X16659955] [Medline: 27450571]
27. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
28. Rosso IM, Killgore WDS, Olson EA, Webb CA, Fukunaga R, Auerbach RP, et al. Internet-based cognitive behavior therapy for major depressive disorder: a randomized controlled trial. Depress Anxiety 2017 Mar;34(3):236-245 [FREE Full text] [doi: 10.1002/da.22590] [Medline: 28009467]
29. Sandoval LR, Buckey JC, Ainslie R, Tombari M, Stone W, Hegel MT. Randomized controlled trial of a computerized interactive media-based problem solving treatment for depression. Behav Ther 2017 May;48(3):413-425 [FREE Full text] [doi: 10.1016/j.beth.2016.04.001] [Medline: 28390503]
30. Schure MB, Lindow JC, Greist JH, Nakonezny PA, Bailey SJ, Bryan WL, et al. Use of a fully automated internet-based cognitive behavior therapy intervention in a community population of adults with depression symptoms: randomized controlled trial. J Med Internet Res 2019 Nov 18;21(11):e14754 [FREE Full text] [doi: 10.2196/14754] [Medline: 31738173]
31. Stiles-Shields C, Montague E, Kwasny MJ, Mohr DC. Behavioral and cognitive intervention strategies delivered via coached apps for depression: pilot trial. Psychol Serv 2019 May;16(2):233-238. [doi: 10.1037/ser0000261] [Medline: 30407055]
32. Tomasino KN, Lattie EG, Ho J, Palac HL, Kaiser SM, Mohr DC. Harnessing peer support in an online intervention for older adults with depression. Am J Geriatr Psychiatry 2017 Oct;25(10):1109-1119. [doi: 10.1016/j.jagp.2017.04.015] [Medline: 28571785]
33. Geraedts AS, Kleiboer AM, Twisk J, Wiezer NM, van Mechelen W, Cuijpers P. Long-term results of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014;16(7):e168 [FREE Full text] [doi: 10.2196/jmir.3539] [Medline: 25008127]
34. Geraedts AS, van Dongen JM, Kleiboer AM, Wiezer NM, van Mechelen W, Cuijpers P, et al. Economic evaluation of a web-based guided self-help intervention for employees with depressive symptoms: results of a randomized controlled trial. J Occup Environ Med 2015 Jun;57(6):666-675. [doi: 10.1097/JOM.0000000000000423] [Medline: 25741795]

35. Kelders SM, Bohlmeijer ET, Pots WTM, van Gemert-Pijnen EWC. Comparing human and automated support for depression: fractional factorial randomized controlled trial. Behav Res Ther 2015 Sep;72:72-80. [doi: 10.1016/j.brat.2015.06.014] [Medline: 26196078]
36. Kenter RMF, Cuijpers P, Beekman A, van Straten A. Effectiveness of a web-based guided self-help intervention for outpatients with a depressive disorder: short-term results from a randomized controlled trial. J Med Internet Res 2016 Mar 31;18(3):e80 [FREE Full text] [doi: 10.2196/jmir.4861] [Medline: 27032449]
37. Kleiboer A, Donker T, Seekles W, van Straten A, Riper H, Cuijpers P. A randomized controlled trial on the role of support in Internet-based problem solving therapy for depression and anxiety. Behav Res Ther 2015 Sep;72:63-71. [doi: 10.1016/j.brat.2015.06.013] [Medline: 26188373]
38. Pots WTM, Fledderus M, Meulenbeek PAM, ten Klooster PM, Schreurs KMG, Bohlmeijer ET. Acceptance and commitment therapy as a web-based intervention for depressive symptoms: randomised controlled trial. Br J Psychiatry 2016 Jan;208(1):69-77. [doi: 10.1192/bjp.bp.114.146068] [Medline: 26250745]
39. Beiwinkel T, Eißing T, Telle N, Siegmund-Schultze E, Rössler W. Effectiveness of a web-based intervention in reducing depression and sickness absence: randomized controlled trial. J Med Internet Res 2017 Jun 15;19(6):e213 [FREE Full text] [doi: 10.2196/jmir.6546] [Medline: 28619701]
40. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: results of a pragmatic randomized controlled trial. J Affect Disord 2018 Feb;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]
41. Bücker L, Schnakenberg P, Karyotaki E, Moritz S, Westermann S. Diminishing effects after recurrent use of self-guided internet-based interventions in depression: randomized controlled trial. J Med Internet Res 2019 Oct 02;21(10):e14240 [FREE Full text] [doi: 10.2196/14240] [Medline: 31579014]
42. Ebert DD, Buntrock C, Lehr D, Smit F, Riper H, Baumeister H, et al. Effectiveness of web- and mobile-based treatment of subthreshold depression with adherence-focused guidance: a single-blind randomized controlled trial. Behav Ther 2018 Dec;49(1):71-83. [doi: 10.1016/j.beth.2017.05.004] [Medline: 29405923]
43. Gräfe V, Berger T, Hautzinger M, Hohagen F, Lutz W, Meyer B, et al. Health economic evaluation of a web-based intervention for depression: the EVIDENT-trial, a randomized controlled study. Health Econ Rev 2019 Jun 07;9(1):16 [FREE Full text] [doi: 10.1186/s13561-019-0233-y] [Medline: 31175475]
44. Gräfe V, Moritz S, Greiner W. Health economic evaluation of an internet intervention for depression (deprexis), a randomized controlled trial. Health Econ Rev 2020 Jun 16;10(1):19 [FREE Full text] [doi: 10.1186/s13561-020-00273-0] [Medline: 32548724]
45. Jelinek L, Arlt S, Moritz S, Schröder J, Westermann S, Cludius B. Brief web-based intervention for depression: randomized controlled trial on behavioral activation. J Med Internet Res 2020 Mar 26;22(3):e15312 [FREE Full text] [doi: 10.2196/15312] [Medline: 32213470]
46. Löbner M, Dorow M, Pabst A, Stein J, Kersting A, Luppa M, et al. [Online self-management of depression: dose-response-relationship within a primary care setting]. Fortschr Neurol Psychiatr 2019 Mar;87(3):181-186. [doi: 10.1055/a-0849-9802] [Medline: 30891721]
47. Lüdtke T, Pult LK, Schröder J, Moritz S, Bücker L. A randomized controlled trial on a smartphone self-help application (Be Good to Yourself) to reduce depressive symptoms. Psychiatry Res 2018 Nov;269:753-762. [doi: 10.1016/j.psychres.2018.08.113] [Medline: 30273901]
48. Zagorscak P, Heinrich M, Bohn J, Stein J, Knaevelsrud C. How individuals change during internet-based interventions for depression: a randomized controlled trial comparing standardized and individualized feedback. Brain Behav 2020 Jan;10(1):e01484 [FREE Full text] [doi: 10.1002/brb3.1484] [Medline: 31777204]
49. Zwerenz R, Becker J, Johansson R, Frederick RJ, Andersson G, Beutel ME. Transdiagnostic, psychodynamic web-based self-help intervention following inpatient psychotherapy: results of a feasibility study and randomized controlled trial. JMIR Ment Health 2017 Oct 16;4(4):e41 [FREE Full text] [doi: 10.2196/mental.7889] [Medline: 29038094]
50. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. Behav Res Ther 2018 Oct;109:75-83. [doi: 10.1016/j.brat.2018.08.003] [Medline: 30125790]
51. Clarke J, Proudfoot J, Birch M, Whitton AE, Parker G, Manicavasagar V, et al. Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: secondary analysis of a randomised controlled trial. BMC Psychiatry 2014;14:272 [FREE Full text] [doi: 10.1186/s12888-014-0272-1] [Medline: 25252853]
52. Crisp D, Griffiths K, Mackinnon A, Bennett K, Christensen H. An online intervention for reducing depressive symptoms: secondary benefits for self-esteem, empowerment and quality of life. Psychiatry Res 2014 Apr 30;216(1):60-66. [doi: 10.1016/j.psychres.2014.01.041] [Medline: 24534125]
53. Fogarty AS, Proudfoot J, Whittle EL, Clarke J, Player MJ, Christensen H, et al. Preliminary evaluation of a brief web and mobile phone intervention for men with depression: men's positive coping strategies and associated depression, resilience, and work and social functioning. JMIR Ment Health 2017 Aug 10;4(3):e33 [FREE Full text] [doi: 10.2196/mental.7769] [Medline: 28798009]

54. Pictet A, Jermann F, Ceschi G. When less could be more: investigating the effects of a brief internet-based imagery cognitive bias modification intervention in depression. Behav Res Ther 2016 Dec;84:45-51. [doi: 10.1016/j.brat.2016.07.008] [Medline: 27485252]
55. Wagner B, Horn AB, Maercker A. Internet-based versus face-to-face cognitive-behavioral intervention for depression: a randomized controlled non-inferiority trial. J Affect Disord 2014 Jan;152-154:113-121. [doi: 10.1016/j.jad.2013.06.032] [Medline: 23886401]
56. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
57. Castro A, García-Palacios A, López-Del-Hoyo Y, Mayoral F, Pérez-Ara, Baños RM, et al. Predictors of adherence in three low-intensity intervention programs applied by ICTs for depression in primary care. Int J Environ Res Public Health 2021 Feb 11;18(4):246-261 [FREE Full text] [doi: 10.3390/ijerph18041774] [Medline: 33670353]
58. Gili M, Castro A, García-Palacios A, Garcia-Campayo J, Mayoral-Cleries F, Botella C, et al. Efficacy of three low-intensity, internet-based psychological interventions for the treatment of depression in primary care: randomized controlled trial. J Med Internet Res 2020 Jun 05;22(6):e15845 [FREE Full text] [doi: 10.2196/15845] [Medline: 32501276]
59. Mira A, Bretón-López J, García-Palacios A, Quero S, Baños RM, Botella C. An Internet-based program for depressive symptoms using human and automated support: a randomized controlled trial. Neuropsychiatr Dis Treat 2017;13:987-1006 [FREE Full text] [doi: 10.2147/NDT.S130994] [Medline: 28408833]
60. Mira A, Bretón-López J, Enrique A, Castilla D, García-Palacios A, Baños R, et al. Exploring the incorporation of a positive psychology component in a cognitive behavioral internet-based program for depressive symptoms. results throughout the intervention process. Front Psychol 2018;9:2360 [FREE Full text] [doi: 10.3389/fpsyg.2018.02360] [Medline: 30555384]
61. Agyapong VIO, Juhás M, Ohinmaa A, Omeje J, Mrklas K, Suen VYM, et al. Randomized controlled pilot trial of supportive text messages for patients with depression. BMC Psychiatry 2017 Aug 02;17(1):286 [FREE Full text] [doi: 10.1186/s12888-017-1448-2] [Medline: 28768493]
62. Khatri N, Marziali E, Tchernikov I, Shepherd N. Comparing telehealth-based and clinic-based group cognitive behavioral therapy for adults with depression and anxiety: a pilot study. Clin Interv Aging 2014;9:765-770 [FREE Full text] [doi: 10.2147/CIA.S57832] [Medline: 24855345]
63. Lappalainen P, Granlund A, Siltanen S, Ahonen S, Vitikainen M, Tolvanen A, et al. ACT Internet-based vs face-to-face? A randomized controlled trial of two ways to deliver Acceptance and Commitment Therapy for depressive symptoms: an 18-month follow-up. Behav Res Ther 2014 Oct;61:43-54. [doi: 10.1016/j.brat.2014.07.006] [Medline: 25127179]
64. Lappalainen P, Langrial S, Oinas-Kukkonen H, Tolvanen A, Lappalainen R. Web-based acceptance and commitment therapy for depressive symptoms with minimal support: a randomized controlled trial. Behav Modif 2015 Aug 6;39(6):805-834. [doi: 10.1177/0145445515598142] [Medline: 26253644]
65. Flygare A, Engström I, Hasselgren M, Jansson-Fröjmark M, Frejgrim R, Andersson G, et al. Internet-based CBT for patients with depressive disorders in primary and psychiatric care: is it effective and does comorbidity affect outcome? Internet Interv 2020 Mar;19:100303 [FREE Full text] [doi: 10.1016/j.invent.2019.100303] [Medline: 32055451]
66. Kivi M, Eriksson MCM, Hange D, Petersson E, Vernmark K, Johansson B, et al. Internet-based therapy for mild to moderate depression in Swedish primary care: short term results from the PRIM-NET randomized controlled trial. Cogn Behav Ther 2014;43(4):289-298 [FREE Full text] [doi: 10.1080/16506073.2014.921834] [Medline: 24911260]
67. Blackwell SE, Browning M, Mathews A, Pictet A, Welch J, Davies J, et al. Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: a randomized controlled trial. Clin Psychol Sci 2015 Jan;3(1):91-111 [FREE Full text] [doi: 10.1177/2167702614560746] [Medline: 25984421]
68. Lambert JD, Greaves CJ, Farrand P, Price L, Haase AM, Taylor AH. Web-based intervention using behavioral activation and physical activity for adults with depression (The eMotion Study): pilot randomized controlled trial. J Med Internet Res 2018 Jul 16;20(7):e10112 [FREE Full text] [doi: 10.2196/10112] [Medline: 30012547]
69. Hatcher S, Whittaker R, Patton M, Miles WS, Ralph N, Kercher K, et al. Web-based therapy plus support by a coach in depressed patients referred to secondary mental health care: randomized controlled trial. JMIR Ment Health 2018 Jan 23;5(1):e5 [FREE Full text] [doi: 10.2196/mental.8510] [Medline: 29362207]
70. Schuster R, Fichtenbauer I, Sparr VM, Berger T, Laireiter A. Feasibility of a blended group treatment (bGT) for major depression: uncontrolled interventional study in a university setting. BMJ Open 2018 Mar 12;8(3):e018412 [FREE Full text] [doi: 10.1136/bmjopen-2017-018412] [Medline: 29530905]
71. 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 [FREE Full text] [doi: 10.1016/j.brat.2015.10.005] [Medline: 26523885]
72. 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] [Medline: 27280960]

73. Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med 2013;10(5):e1001454 [FREE Full text] [doi: 10.1371/journal.pmed.1001454] [Medline: 23723742]
74. Ebert DD, Berking M, Cuijpers P, Lehr D, Pörtner M, Baumeister H. Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms: a randomized controlled trial. J Affect Disord 2015 May 1;176:9-17. [doi: 10.1016/j.jad.2015.01.056] [Medline: 25682378]
75. 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] [Medline: 20183695]
76. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
77. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-869. [doi: 10.1586/ern.12.63] [Medline: 22853793]
78. Wright JH, Mishkind M, Eells TD, Chan SR. Computer-assisted cognitive-behavior therapy and mobile apps for depression and anxiety. Curr Psychiatry Rep 2019 Jun 27;21(7):62. [doi: 10.1007/s11920-019-1031-2] [Medline: 31250242]
79. Shen N, Levitan M, Johnson A, Bender JL, Hamilton-Page M, Jadad AAR, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
80. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]

Abbreviations

ACT: acceptance and commitment therapy
CBT: cognitive behavioral therapy
CG: control group
DHI: digital health intervention
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EU: European Union
IG: intervention group
LZG.NRW: Landeszentrum Gesundheit North Rhine-Westphalia
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses
RCT: randomized controlled trial
SCID: Structured Clinical Interview for DSM-5
TAU: treatment as usual

Edited by J Torous; submitted 04.12.20; peer-reviewed by S Mathews, J D'Arcey; comments to author 14.12.20; revised version received 15.01.21; accepted 19.02.21; published 16.04.21.

Please cite as:
Tokgöz P, Hrynyschyn R, Hafner J, Schönfeld S, Dockweiler C
Digital Health Interventions in Prevention, Relapse, and Therapy of Mild and Moderate Depression: Scoping Review
JMIR Ment Health 2021;8(4):e26268
URL: https://mental.jmir.org/2021/4/e26268
doi:10.2196/26268
PMID:33861201

Review

Evidence on Digital Mental Health Interventions for Adolescents and Young People: Systematic Overview

Susanna Lehtimaki[1*], MPH, MSSc; Jana Martic[1], MSc; Brian Wahl[1,2], MPH, PhD; Katherine T Foster[3,4], PhD; Nina Schwalbe[1,5,6*], MPH

[1]Spark Street Advisors, New York, NY, United States

[2]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]Department of Psychology, University of Washington, Seattle, WA, United States

[4]Department of Global Health, University of Washington, Seattle, WA, United States

[5]Heilbrunn Department of Population and Family Health, Columbia Mailman School of Public Health, New York, NY, United States

[6]United Nations University International Institute for Global Health, Kuala Lumpur, Malaysia

[*]these authors contributed equally

Corresponding Author:
Nina Schwalbe, MPH
Heilbrunn Department of Population and Family Health
Columbia Mailman School of Public Health
722 W 168th Street
New York, NY, 10032
United States
Phone: 1 9172262645
Email: nschwalbe@ssc.nyc

Abstract

Background: An estimated 1 in 5 adolescents experience a mental health disorder each year; yet because of barriers to accessing and seeking care, most remain undiagnosed and untreated. Furthermore, the early emergence of psychopathology contributes to a lifelong course of challenges across a broad set of functional domains, so addressing this early in the life course is essential. With increasing digital connectivity, including in low- and middle-income countries, digital health technologies are considered promising for addressing mental health among adolescents and young people. In recent years, a growing number of digital health interventions, including more than 2 million web-based mental health apps, have been developed to address a range of mental health issues.

Objective: This review aims to synthesize the current evidence on digital health interventions targeting adolescents and young people with mental health conditions, aged between 10-24 years, with a focus on effectiveness, cost-effectiveness, and generalizability to low-resource settings (eg, low- and middle-income countries).

Methods: We searched MEDLINE, PubMed, PsycINFO, and Cochrane databases between January 2010 and June 2020 for systematic reviews and meta-analyses on digital mental health interventions targeting adolescents and young people aged between 10-24 years. Two authors independently screened the studies, extracted data, and assessed the quality of the reviews.

Results: In this systematic overview, we included 18 systematic reviews and meta-analyses. We found evidence on the effectiveness of computerized cognitive behavioral therapy on anxiety and depression, whereas the effectiveness of other digital mental health interventions remains inconclusive. Interventions with an in-person element with a professional, peer, or parent were associated with greater effectiveness, adherence, and lower dropout than fully automatized or self-administered interventions. Despite the proposed utility of digital interventions for increasing accessibility of treatment across settings, no study has reported sample-specific metrics of social context (eg, socioeconomic background) or focused on low-resource settings.

Conclusions: Although digital interventions for mental health can be effective for both supplementing and supplanting traditional mental health treatment, only a small proportion of existing digital platforms are evidence based. Furthermore, their cost-effectiveness and effectiveness, including in low- and middle-income countries, have been understudied. Widespread adoption and scale-up of digital mental health interventions, especially in settings with limited resources for health, will require more rigorous and consistent demonstrations of effectiveness and cost-effectiveness vis-à-vis the type of service provided, target population, and the current standard of care.

(JMIR Ment Health 2021;8(4):e25847) doi:10.2196/25847

KEYWORDS

digital health; adolescent health; young people; mental health; digital technologies

Introduction

Background

Mental health issues remain underdiagnosed and undertreated among adolescents and young people (aged 10-24 years) [1]. Ignored by many health and social services and policies worldwide [2], adolescents and young people are particularly vulnerable to many conditions affecting mental health. Nearly 50% of mental health disorders begin by the age of 14 years, and 75% of mental health disorders begin by the age of 24 years [3]; an estimated 1 in 5 adolescents experience a mental health disorder each year [4]. The emergence of symptom sequelae, even below the diagnostic threshold, signals an increased vulnerability to life course–persistent mental health problems and consequences if not addressed early. Among men and women aged between 15-19 years, suicide, which is more common among young people than adults [5], is one of the top 3 causes of death worldwide, and depression is among the leading causes of disability for those aged between 10-19 years [6].

At the same time, young people are growing up in the digital world and accessing the internet at increasingly younger ages [7]. As the most connected age group in the population, more than 70% of young people aged between 15-24 years are "online" [8]. Although there are income-based and geographical disparities in digital access, 43% of people in low- and middle-income countries use the internet, and even in low-income countries, 72% of people have access to mobile phones, and 16% of people have access to the internet [9].

Although there are clearly some negative effects of technology on this age group, including behavioral addiction, cyber-bullying, depression, sexual exploitation, and abuse [10-12], the use of digitally enabled technology is considered a promising platform for preventing morbidity and enhancing well-being and quality of life [13]. Critically, digital technologies may offer especially critical support for adolescents and young people in low-resource settings where barriers to care may be numerous and insurmountable.

Given the increasing number of adolescents and young people using digital technologies, digital mental health interventions are considered to have the specific potential to support mental health and well-being in this group [14,15]. Specifically, digital technology could provide opportunities to access mental health services and information while also increasing patient empowerment, participation [16], and help-seeking and helping to overcome the stigma that is often linked to mental health services [17]. With more than 2 million mental health apps already available, including 40,000 classified as medical [18], the demand for this innovation is evident. However, the plethora of these apps may have outpaced the development of a correspondingly large evidence base on their effectiveness.

Objectives

A number of systematic reviews and meta-analyses have been conducted over the past 10 years on the use of digital technology to enhance mental health among adolescents and young people. A higher-level synthesis of information across these meta-analyses and reviews is needed to identify whether there is converging evidence for their effectiveness and to assess systematic issues with research in this area. Consequently, this systematic overview provides a high-level synthesis of the current evidence on the effectiveness of digital health interventions targeting adolescents and young people (ie, aged 10-24 years as defined by the World Health Organization and others [19,20]; Textbox 1) with diagnosed or self-reported mental health conditions, including affective, behavioral, and trauma-related conditions (eg, anxiety, depression, psychological distress, eating disorders, and posttraumatic stress disorder). Furthermore, it aims to characterize the factors, including digital platforms and design elements used, that contribute to the effectiveness. Finally, it aims to describe the extent to which there is evidence of the economic benefits of such interventions and determine the extent to which previous research in this area may generalize to low-resource settings, including low- and middle-income countries.

The research questions are as follows:

- In adolescents and young people aged between 10 and 24 years, to what extent are digital health interventions effective in addressing mental health conditions, compared with standard face-to-face treatment, placebo, or no treatment?
- What factors contribute to effectiveness (ie, what makes effective interventions effective)?
- To what extent is there evidence on cost-effectiveness?
- To what extent are the findings generalizable to adolescents and young people from a range of settings, including low- and middle-income countries?

Textbox 1. Definitions of key terms.

Adolescents and young people

- According to the World Health Organization, adolescents are individuals aged 10-19 years, and young people are individuals aged 10-24 years [19]

Mental health conditions, mental disorders

- Mental health problems with different symptoms, characterized by a combination of abnormal thoughts, perceptions, emotions, behavior, and relationships with others [21,22]

Digital mental health intervention

- Information, support, and therapy for mental health conditions delivered through an electronic medium with the aim of treating, alleviating, or managing symptoms [23,24]

Cognitive behavioral therapy (CBT), computerized CBT (cCBT)

- A form of psychological treatment to identify maladaptive patterns of thinking, emotional response, or behavior and substituting them with desirable patterns [25]. cCBT refers to computerized implementation of CBT

Effectiveness, effect

- The ability of an intervention to produce intended outcomes, estimated by comparing the intervention with no intervention (ie, better than nonactive control) and/or an existing evidence-based intervention (ie, no difference from active control) [26]

Active control

- A comparison group receiving standard treatment, including face-to-face therapy, alternative therapy, or materials [26]

Nonactive control

- A comparison group not receiving or performing any activity. These may include placebo treatment, no treatment, or assigned to a waitlist to receive intervention after completion of the trial [26]

Methods

Search Strategy and Selection Criteria

The review was conducted using a predefined protocol. We conducted an electronic review of the literature from the MEDLINE, PubMed, PsycINFO, and Cochrane databases. The review was limited to peer-reviewed articles published in English between January 1, 2011, and July 6, 2020. We used a combination of keywords: (“digital,” “mHealth,” “eHealth,” “web-based,” “internet-based,” “mobile phone,” “text message,” “SMS,” “artificial intelligence”) AND (“adolescen*,” “youth,” “young,” “child,” “student”) AND (“mental health,” “wellbeing”). Our search was limited to overview types of studies, such as meta-analyses and systematic reviews.

Identified references were screened independently by 2 reviewers (SL and JM) by conducting an abstract and title search with the following inclusion criteria, following a predefined PICOS (Population, Intervention, Comparator, Outcome, Setting) framework:

- Population: Adolescents and young people, defined as primarily aged between 10 and 24 years (or if older participants were included, the mean age was <25 years), with a mental health condition, including anxiety, affective, and behavioral conditions (diagnosed and self-reported)
- Intervention: Consumer-facing, partially or fully self-administered, mental health intervention delivered through a digital platform (eg, web-based, computer, or mobile phone)
- Comparator: Active (ie, standard nondigital care and alternative materials) or passive control (ie, placebo and no treatment)
- Outcome: Mental health improvement as reported by studies (ie, diagnosed or self-reported mental health conditions, including affective, behavioral, and trauma-related conditions)
- Setting: Nonclinical, nonfacility-based setting in any country

Potentially relevant studies identified through the screening process were assessed independently for final inclusion by 2 reviewers (SL and JM) after being acquired in full text. References were excluded if they were not exclusive to this age group; were delivered at the health care facility (eg, telemedicine by clinicians); targeted adolescents and young people with chronic diseases, such as HIV, diabetes, or cancer; targeted adolescents and young people with mental and behavioral disorders because of psychoactive substance use; or were primarily addressing parenting skills or targeting parents. Study protocols and nonpeer-reviewed papers were excluded.

Data Extraction and Quality Assessment

In total, 2 reviewers (SL and JM) independently extracted information from the studies, building a matrix including data on participants (age and other available background characteristics), interventions, mental health issues addressed, setting (eg, delivery platforms and countries), and key findings in terms of clinical effectiveness. The reviewers also assessed the quality of the articles by using the AMSTAR 2 (A

Measurement Tool to Assess Systematic Reviews) [27] tool, which is a validated tool to analyze the quality of systematic reviews and meta-analyses with ratings from high to critically low. The guidance document of the tool [27] was thoroughly followed. Any disagreement in either of these actions was resolved through discussion.

Data Synthesis

We synthesized evidence from the articles describing the effectiveness of digital mental health interventions against clinical outcomes, therapy used, and digital platform deployed as well as reviewed factors associated with effectiveness, sustainability of outcomes, completion, and adherence. Finally, we reviewed and synthesized the extent to which there was evidence on the cost-effectiveness and the potential generalizability of the findings to low- and middle-income countries. Given the high heterogeneity of the studies, we did not conduct a statistical analysis.

Results

Overview

The initial search yielded 1295 articles. After excluding duplicate references, the number of articles was reduced to 1098. The search strategy was complemented by a manual search of reference lists of key articles, which yielded an additional 8 articles for eligibility assessment (Figure 1).

After title screening, we conducted full-text appraisal and excluded articles that did not meet the inclusion criteria. A total of 18 articles were finally included (Table 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.

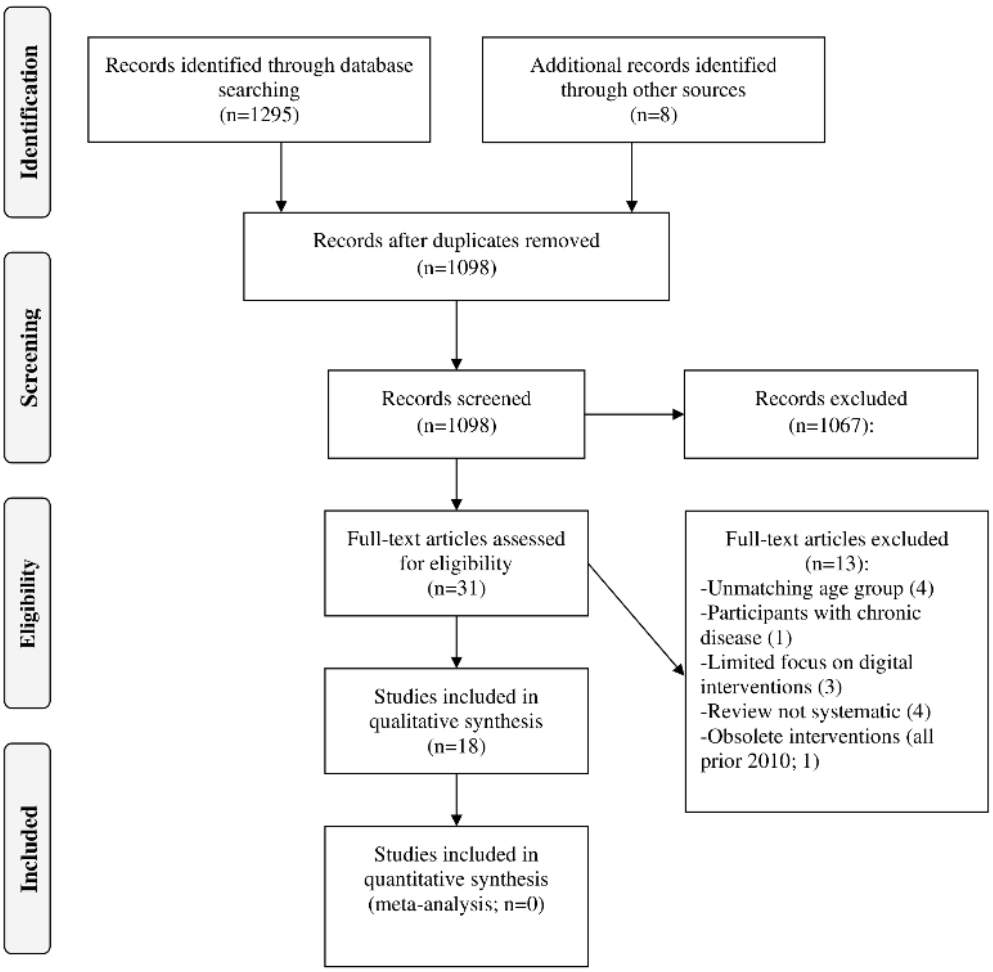


Table 1. Study characteristics.

Author	Primary studies	Intervention, digital platform	Outcome	Comparison interventions	Total sample	Participants age (range and mean, if available)	Geographical coverage
Barnes et al (2018) [28]	Systematic review of 2 RCTs[a] and 3 NRSs[b] on 3 games	Video games	Anxiety	Nonactive (alternative nontherapeutic video game)	410	Up to 19 years	China, Hong Kong, the Netherlands, and the United Kingdom
Bry et al (2018) [29]	Systematic review and content analysis of 121 apps	Mobile apps	Anxiety	No control group	Not reported	Not defined (children and adolescents)	Not defined
Grist et al (2017) [30]	Systematic review of 23 NRTs[c] and 1 RCT on 15 apps	Mobile apps	Mental health, well-being, anxiety, depression, suicide, obsessive-compulsive disorder, and eating disorders	Active (nondigital intervention; nb. Only one RCT included in the review)	1054	Up to 18 years	Australia, Canada, Denmark, Ireland, the Netherlands, and the United States
Hollis et al (2017) [24]	Systematic review of 30 RCTs; meta-review of 21 articles on 147 interventions	Internet-based interventions, mobile apps, and eHealth	Anxiety, depression, attention deficit hyperactivity disorder, autism spectrum disorder, psychosis, eating disorders, and posttraumatic stress disorder	Mixed nonactive (waitlist, no intervention) and active (attention control group, limited intervention)	5333	Up to 25 years	Australia, China, the Netherlands, New Zealand, Norway, Israel, Sweden, Switzerland, the United Kingdom, and the United States
Davies et al (2014) [31]	Systematic review of 17 RCTs and meta-analysis of 14 RCTs	Computer-delivered or web-based interventions	Anxiety, depression, psychological distress, and stress	Mixed nonactive (no treatment, waitlist) and active (alternative materials)	1480	17-51 years; mean 22.6 years	Australia, Canada, Norway, Spain, the United Kingdom, and the United States
Farrer et al (2013) [32]	Systematic review of 26 RCTs and 1 randomized trial	Internet-based, audio, virtual reality, and computer programs	Anxiety and depression	Mixed nonactive (no intervention, waitlist) and active (attention control group)	Not reported	18-25 years	Australia, Belgium, China, Italy, the Netherlands, Spain, the United Kingdom, and the United States
Valimaki et al (2017) [33]	Systematic review of 22 RCTs and meta-analysis of 15 RCTs	Internet-based interventions	Depression, anxiety, and stress	Mixed nonactive and active (not specified)	4979	10-24 years	Australia, Canada, China, the Netherlands, New Zealand, Norway, the United Kingdom, and the United States
Harrer et al (2019) [34]	Systematic review of 48 randomized trials	Internet-based psychological interventions	Anxiety, depression, stress, sleep problems, eating disorders, and well - being	Mixed nonactive (waitlist, placebo) and active (diaries, recommendations for behavior change)	10,583	Up to 29 years; mean 22 years	Australia, Canada, Finland, Germany, Ireland, Norway, Romania, Spain, Sweden, the United Kingdom, and the United States

Author	Primary studies	Intervention, digital platform	Outcome	Comparison interventions	Total sample	Participants age (range and mean, if available)	Geographical coverage
Garrido et al (2019) [35]	Systematic review of 27 RCTs and 13 NRTs on 32 interventions and meta-analysis of 15 RCTs	Computer, web-based, and smartphone-delivered intervention	Anxiety and depression	Mixed nonactive (waitlist) and active (alternative therapeutic intervention)	16,874	12-25 years	Australia, Canada, Chile, China, Hong Kong, Ireland, Japan, New Zealand, Northern Europe, the United Kingdom, and the United States
Pretorius et al (2019) [36]	Systematic review of 27 qualitative, feasibility, and comparative studies and 1 RCT	Web-based help-seeking interventions	Psychological distress	No control group	Not reported	12-25 years	Australia, Canada, Ireland, the Netherlands, the United Kingdom, and the United States
Ridout et al (2018) [37]	Systematic review of 9 descriptive studies on 5 interventions	Social networking sites	Depression, psychosis, health literacy, social support, and general well-being	No control group	Not reported	Up to 25 years	Australia, China, Hong Kong, and the United States
Podina et al (2016) [38]	Meta-analysis of 8 RCTs	cCBT[d]	Anxiety	Mixed nonactive (waitlist) and active (standard CBT[e])	404	7-18 years	Australia, Canada, Spain, and the United States
Ebert et al (2015) [39]	Meta-analysis of 13 RCTs	cCBT	Anxiety and depression	Nonactive (no treatment, placebo)	796	Up to 25 years	Australia, the Netherlands, New Zealand, Sweden, the United Kingdom, and the United States
Pennant et al (2015) [40]	Systematic review and meta-analysis of 27 RCTs	cCBT	Anxiety and depression	Mixed nonactive (waitlist, placebo, no intervention) and active (standard CBT)	3389	5-25 years	Australia, China, the Netherlands, New Zealand, Israel, Sweden, the United Kingdom, and the United States
Ye et al (2014) [41]	Meta-analysis of 7 RCTs	cCBT and SMS	Anxiety and depression	Mixed nonactive (waitlist) and active (standard CBT, alternative intervention)	569	7- 25 years	Australia and the United States
Grist et al (2019) [42]	Meta-analysis of 34 RCTs on 29 interventions	cCBT, computer-delivered attention, or cognitive bias modification programs	Anxiety and depression	Mixed nonactive (waitlist, placebo) and active (face-to-face or alternative therapeutic interventions)	3113	Up to 18 years	Australia, Canada, China, Ireland, Israel, the Netherlands, New Zealand, Sweden, Thailand, the United Kingdom, and the United States
Vigerland et al (2016) [43]	Meta-analysis of 24 RCTs	cCBT	Multiple psychiatric and psychosomatic conditions	Mixed nonactive (waitlist) and active (standard CBT, alternative intervention)	1882	Up to 18 years	Australia, Canada, Germany, the Netherlands, Sweden, and the United States

Author	Primary studies	Intervention, digital platform	Outcome	Comparison interventions	Total sample	Participants age (range and mean, if available)	Geographical coverage
Clarke et al (2015) [44]	Systematic review of 14 RCTs and 14 NRSs on 21 interventions	Mixed web-based interventions and more than half (8/15) cCBT	Mental health promotion and prevention	Mixed nonactive (waitlist, placebo, no intervention) and active (limited intervention)	10,779	12-25 years	Australia, Canada, China, Germany, Ireland, Israel, the Netherlands, Norway, and the United States

[a]RCT: randomized controlled trial.

[b]NRS: nonrandomized study.

[c]NRT: nonrandomized trial.

[d]cCBT: computerized cognitive behavioral therapy.

[e]CBT: cognitive behavioral therapy.

Effectiveness Across Clinical Symptom Targets

In terms of clinical outcomes, most systematic reviews and meta-analyses included in this review focused on anxiety (n=4), depression (n=3), anxiety and depression together (n=11), or anxiety and depression with stress (n=3). To a lesser degree, analyses focused on general well-being (n=4). In addition, eating disorders (n=2), psychosis (n=2), attention-deficit/hyperactivity disorder (ADHD; n=1), autism spectrum disorder (n=1), sleep problems (n=1), suicide prevention (n=1), obsessive-compulsive disorder (n=1), role functioning (n=1), phobias (n=1), and posttraumatic stress disorder (n=1) were clinical outcomes explored in the reviews.

Evidence on the benefits of digital mental health interventions was found for anxiety, depression, and stress when compared with nonactive controls, defined primarily as groups to which no treatment was provided or on those put on a waitlist for services. However, compared with active controls, defined as those undergoing or receiving some type of treatment, they appear to be similarly effective (Table 2).

A meta-analysis by Harrer et al [34] on web-based interventions mostly delivered through a dedicated website found small effects on depression (Hedges *g*=0.18; 95% CI 0.08-0.27), anxiety (Hedges *g*=0.27; 95% CI 0.13-0.40), and stress (Hedges *g*=0.20; 95% CI 0.02-0.38) compared with nonactive controls consisting of waitlist or placebo control groups.

A meta-analysis by Davies et al [31] on mixed web-based and computer-delivered interventions for depression, anxiety, and stress found a small effect of digital interventions in comparison with active controls that received alternative materials (for anxiety, pooled standardized mean difference [SMD] −0.18; 95% CI −0.98 to 0.62; *P*=.66 and for depression, pooled SMD −0.28; 95% CI −0.75 to −0.20; *P*=.25), whereas a medium effect was found when compared with nonactive controls. When compared with a nonactive control, there was some effect of decreasing anxiety (pooled SMD −0.56; 95% CI −0.77 to −0.35; *P*<.001), depression (pooled SMD −0.43; 95% CI −0.63 to −0.22; *P*<.001), and stress (pooled SMD −0.73; 95% CI −1.27 to −0.19; *P*=.008).

A meta-analysis by Garrido et al [35] that focused on depression found a small pooled effect size of digital mental health interventions in comparison with nonactive controls (Cohen *d*=0.33; 95% CI 0.11-0.55), whereas the pooled effect size of studies comparing an intervention group with active controls, mostly receiving alternative materials, including website content, showed no significant differences (Cohen *d*=0.14; 95% CI −0.04 to 0.31).

A systematic review by Farrer et al [32] exploring 51 digital interventions using different delivery methods addressing mostly depression, anxiety, and stress found that nearly half of the interventions (24/51, 47%) were associated with at least one positive outcome after the intervention compared with the control group (nonactive and attention controls) and nearly one-third of the interventions (15/51, 29%) failed to report a significant effect. For interventions targeting both symptoms of depression and anxiety (n=8), in comparison with mixed control groups (nonactive and active), effect sizes ranged significantly from −0.07 to 3.04 (overall median 0.54; [effect size] targeting depression symptoms=0.48 and targeting anxiety symptoms=0.77). For interventions targeting only anxiety (n=10), effect sizes ranged from 0.07 to 2.66 (median 0.84). However, the authors of these reviews could not calculate effect sizes for almost two-thirds of the interventions (33/51, 64%) because of insufficient or unavailable meta-data across the reviewed studies [32].

Outcomes of interventions for ADHD, autism spectrum disorders, eating disorders, psychosis, and posttraumatic stress were reported in 3 systematic reviews and one meta-analysis [24,32,34]. Hollis et al [24] demonstrated inconsistent results on the effectiveness of digital interventions for ADHD, autism, psychosis, or eating disorders, limited by the small number of studies and the high degree of variability in reliance on evidence-based treatments. Farrer et al [32] demonstrated the effectiveness of virtual reality or video exposure interventions on arachnophobia or acrophobia. In addition, Harrer et al [34] found moderate effects on eating disorder symptoms (Hedges *g*=0.52; 95% CI 0.22-0.83) and role functioning (Hedges *g*=0.41; 95% CI 0.26-0.56) in comparison with active and nonactive controls (predominantly waitlist control) but no effect on general well-being in comparison with placebo intervention (Hedges *g*=0.15; 95% CI −0.20 to 0.50).

Table 2. Key findings of the included studies.

Reference	Effectiveness	Contributing factors	Cost-effectiveness	Inclusion of data on low- and middle-income countries	Quality of included studies	Quality of review (AMSTAR[a])
Barnes et al (2018) [28]	Although early findings suggest that therapeutic games have the potential to lead to clinically measurable reductions in symptoms in adolescents with anxiety, evidence on the effectiveness is extremely limited. On the basis of 2 RCTs[b] included in this review, no difference in anxiety outcomes is found between the intervention and control groups (alternative nontherapeutic videogame).	Not discussed	Not discussed	Limited (China and Hong Kong)	Mean rating of 75% using mixed methods appraisal tool. Only 2 RCTs included in the review.	Critically low
Bry et al (2018) [29]	Evidence-based treatment content within consumer smartphone apps marketed for child and adolescent anxiety is scant, and only a few comprehensive anxiety self-management apps are identified. Half of the sampled apps for anxiety include any evidence-based treatment component, and 23% included two or more evidence-based components.	Not discussed	Low cost but effectiveness unknown	No	N/A[c]	N/A
Grist et al (2017) [30]	Authors conclude that there is currently no evidence to support the effectiveness of apps for adolescents with mental health problems. In 2 RCTs on mobile app for depression, anxiety, and stress, no significant effect is found between intervention (app with self-monitoring) and control (no self-monitoring) groups. Acceptability is generally rated average to high, with adherence ranging from 65% to 83%.	Specific factors: privacy, safety, discretion, and data security; credibility of design and visual appearance; engaging and interactive content; concise, interesting, and trustworthy information; reminders to use; and personalization allowed	Not discussed	No	Issues with quality, including small sample size. Only 2 small RCTs included in the review, both without adequate control group.	Critically low
Hollis et al (2017) [24]	cCBT[d] provides clinical benefits for depression and anxiety when compared with inactive control (waitlist). The benefits for attention deficit hyperactivity disorder and autism are inconsistent, for psychosis are unknown, and eating disorders are no better than waitlist control in regard to symptomology.	Self-guided cCBT has poor uptake and adherence. Human involvement is positively associated with adherence. Adolescents and young people prefer face-to-face over web-based interventions. Specific factors: privacy, safety, discretion, and anonymity; providing concise, interesting, and trustworthy information; and ability to complete interventions on own terms and pace.	Authors note a considerable lack of evidence	Limited (China)	Most studies (18/21) rated as moderate quality, 2 rated as low quality, and 1 rated as high quality using AMSTAR. Methodological issues and high level of heterogeneity in the included studies.	Critically low

Reference	Effectiveness	Contributing factors	Cost-effectiveness	Inclusion of data on low- and middle-income countries	Quality of included studies	Quality of review (AMSTAR[a])
Davies et al (2014) [31]	Web-based and computer-delivered interventions are found effective in improving students' depression (pooled SMD[e] −0.43; 95% CI −0.63 to −0.22; *P*<.001), anxiety (pooled SMD −0.56; 95% CI −0.77 to −0.35; *P*<.001), and stress (pooled SMD −0.73; 95% CI −1.27 to −0.19; *P*=.008) outcomes when compared with inactive controls (no treatment, waitlist). When compared with active controls (alternative materials), no benefits are found for depression, anxiety, and stress.	Not discussed	Not discussed	No	A moderate risk of bias. Quality issues with reporting of methodology, data, and outcome measures. Only 3 studies with active control, with reported skewed data. Heterogeneity of interventions.	Moderate
Farrer et al (2013) [32]	Approximately half (24/51) of the technology-based mental health interventions targeting tertiary students with anxiety or depression are associated with at least one significant positive outcome, and approximately one-third (15/51) fail to find a significant effect. Effect size for interventions targeting symptoms of depression and anxiety range from −0.07 to 3.04 (median 0.54; depression=0.48; anxiety=0.77). Effect size for interventions targeting symptoms of anxiety range from 0.07 to 2.66 (median 0.84). cCBT was the most deployed therapy in 25 of 51 of the interventions.	Not discussed	Included studies do not report cost-effectiveness	Limited (China)	Mean rating 4.42 out of 9 using Cochrane Effective Practice and Organisation of Care Group. Methodological issues with reporting on randomization, intended outcomes, and heterogeneity of interventions. Insufficient data in more than half of the studies (14/27) to calculate effect sizes.	Low
Valimaki et al (2017) [33]	Web-based mental health interventions yield statistically significant effect on depressive (*P*=.02; median 1.68; 95% CI 3.11 to 0.25) and anxiety symptoms (*P*<.001; median 1.47; 95% CI 2.36 to 0.59) when compared with control group (type not specified), but not on stress (*P*=.14; median 1.06; 95% CI 2.44 to 0.33). After 6 months of intervention, significant improvement is found on depressive symptoms (*P*=.01; median 1.78; 95% CI 3.20 to 0.37), on anxiety symptoms (*P*<.001; median 1.47; 95% CI 2.36 to 0.59), and on moods and feelings (*P*=.04; median 5.55; 95% CI 10.88 to 0.22). Dropout of those in intervention groups was higher than those in control groups.	Interventions with human elements, such as face-to-face guidance or telephone follow-ups, are associated with adherence and effect.	Included studies do not assess costs. Authors note a considerable lack of evidence	Limited (China)	Some risk of bias using Review Manager. Issues include biases related to attrition rates, selective reporting, and small sample sizes. Mixed control groups.	High

Reference	Effectiveness	Contributing factors	Cost-effectiveness	Inclusion of data on low- and middle-income countries	Quality of included studies	Quality of review (AMSTAR[a])
Harrer et al (2019) [34]	Internet interventions for university students' mental health have a small effect on anxiety (Hedges *g*=0.27; 95% CI 0.13 to 0.40), depression (Hedges *g*=0.18; 95% CI 0.08 to 0.27), and stress (Hedges *g*=0.20; 95% CI 0.02 to 0.38) when compared with nonactive controls. Moderate effects were found on eating disorder symptoms (Hedges *g*=0.52; 95% CI 0.22 to 0.83) and role functioning (Hedges *g*=0.41; 95% CI 0.26 to 0.56). Effects on well‐being are nonsignificant (Hedges *g*=0.15; 95% CI −0.20 to 0.50).	Guidance does not significantly affect intervention efficacy (*P*≥.05).	Not discussed	Limited (Romania)	Half of the studies with high risk of bias. Moderate to substantial level of heterogeneity and selective reporting.	Low
Garrido et al (2019) [35]	Digital interventions work better than no intervention (Cohen *d*=0.33; 95% CI 0.11 to 0.55) but not better than active alternatives (alternative web-based materials; Cohen *d*=0.14; 95% CI −0.04 to 0.31) in improving depression in young people, when results of different studies are pooled together. Most interventions were based on CBT[f]. Authors conclude that interventions may be clinically significant only if supervised. Engagement and adherence rates are low.	Interventions with supervision have a higher pooled effect size than those without supervision (studies with no intervention controls: Cohen *d*=0.52; 95% CI 0.23 to 0.80 and studies with active controls: Cohen *d*=0.49; 95% CI −0.11 to 1.01). Specific factors: credibility of design and visual appearance; concise, interesting, and trustworthy resources; engaging and interactive tools and content; esthetically attractive; relatable situations, characters, or avatars; and reflect local and cultural differences and needs. Technical glitches as a barrier to complete interventions.	Not discussed	Limited (China, Hong Kong, and Chile)	On the basis of Joanna Brigg Institute appraisal tool and CONSORT (Consolidated Standards of Reporting Trials), 32 of 41 studies with high or unclear overall bias and 9 of 41 with low overall bias.	Low
Pretorius et al (2019) [36]	N/A	Young people value web-based services because of anonymity, accessibility, self-reliance, and ease of use. Theoretical frameworks, including self-determination theory and help-seeking model, should be deployed in research. Specific factors: anonymity, privacy, safety, and discretion; site moderation by professionals; credibility of design and visual appearance; concise, interesting, and trustworthy information; esthetically attractive; flexibility, self-reliance, and control; and 24-h availability.	Not discussed	No	Moderate to strong using Critical Appraisal Skills Program. Heterogeneity of interventions. Only 1 RCT included in the review.	Critically low
Ridout et al (2018) [37]	Social networking sites targeting mental health have significant improvement in mental health knowledge and a number of depressive symptoms in young people, but no improvement in anxiety or psychosis symptoms. The results are not compared with a control group.	Young people value involvement of professionals and peers in social networking sites.	Authors conclude that web-based interventions are cost-effective but provide no evidence	Limited (China and Hong Kong)	No quality assessment performed. On the basis of descriptive studies, no RCTs included in the review.	Critically low

Reference	Effectiveness	Contributing factors	Cost-effectiveness	Inclusion of data on low- and middle-income countries	Quality of included studies	Quality of review (AMSTAR[a])
Podina et al (2016) [38]	cCBT is as effective as standard CBT (Hedges *g*=0.295) and more effective than waitlist (Hedges *g*=1.410) in reducing anxiety symptoms in anxious children and adolescents.	Not discussed	Not discussed	No	No quality assessment performed. No publication bias found. Only 8 RCTs included in the review.	Critically low
Ebert et al (2015) [39]	cCBT for youth is associated with significant moderate to large effects on symptoms of anxiety (Hedges *g*=0.68; 95% CI 0.45 to 0.92; *P*<.001) and depression (Hedges *g*=0.68; 95% CI 0.45 to 0.92; *P*<.001) in comparison with nonactive controls. Effect size on symptoms of anxiety or depression for cCBT was similar to face-to-face CBT (Hedges *g*=0.72 vs Hedges *g*=0.66) and higher than face-to-face CBT targeting depression (Hedges *g*=0.35).	No association between parental involvement and better outcomes (without parental involvement: Hedges *g*=0.83; 95% CI 0.53 to 1.13; *P*<.001; NNT[g]=2.26 and with parental involvement: Hedges *g*=0.64; 95% CI 0.40 to 0.88; *P*<.001; NNT=2.86)	Not discussed	No	Low risk of bias overall. Low heterogeneity	Low
Pennant et al (2015) [40]	cCBT has positive effects for symptoms of anxiety (SMD 0.77; 95% CI 1.45 to 0.09; n=6; number of participants=220) and depression (SMD 0.62; 95% CI 1.13 to 0.11; n=7; number of participants=279) for young people with risk of diagnosed anxiety and depression disorders. cCBT has lower effect size on anxiety (SMD 0.15; 95% CI 0.26 to 0.03; number of participants=1273) and depression (SMD 0.15; 95% CI 0.26 to 0.03; number of participants=1280) in the general population. Evidence for interventions other than cCBT is sparse and inconclusive.	Not discussed	Not discussed	Limited (China)	On the basis of Grading of Recommendations, Assessment, Development and Evaluation evidence quality review, most studies rated from very low (1/17) to low (11/17) to moderate (5/17). Heterogeneity associated with number of outcomes.	Critically low
Ye et al (2014) [41]	When compared with inactive controls, cCBT is effective in reducing anxiety symptoms (SMD −0.52; 95% CI −0.90 to −0.14) but not depression (SMD −0.16; 95% CI −0.44 to 0.12). No significant difference is found when compared with standard face-to-face CBT, suggesting it is as effective.	Not discussed	Included studies do not report on cost-effectiveness	No	On the basis of Quality Assessment Tool for Quantitative Studies, studies rated high (3/7) and moderate (4/7) quality. Only 7 RCTs included in the review.	Critically low

Reference	Effectiveness	Contributing factors	Cost-effectiveness	Inclusion of data on low- and middle-income countries	Quality of included studies	Quality of review (AMSTAR[a])
Grist et al (2019) [42]	A small effect (n=8; Hedges *g*=0.41; 95% CI 0.08 to 0.73; *P*<.01) is found in technology-delivered mental health interventions related to attention bias modification when compared with waitlist controls. Although cCBT interventions yield a medium effect size, attention bias modification programs yield a small effect size, and cognitive bias modification programs yield no effect size.	Therapist support (Cochran Q=27.28; *P*<.001) as well as parental involvement (Cochran Q=24.43; *P*<.001) have a significant effect on effectiveness of and adherence to an intervention. Therapist involvement yields a higher effect size (n=9; Hedges *g*=0.87; 95% CI 0.68 to 1.06; *P*<.001) than predominantly or purely self-administered interventions.	Authors note a considerable lack of evidence	Limited (China)	Most studies rated as low quality and unclear risk using Cochrane Risk of Bias Tool. Most studies (29/34) conducted by program developer. Methodological limitations, small sample size, and nonblinding participants.	Low
Vigerland et al (2016) [43]	cCBT yields moderate effects when compared with waitlist controls (Hedges *g*=0.62; 95% CI 0.41 to 0.84).	Not discussed	Authors note a considerable lack of evidence	No	Quality varied largely across the studies; Moncrieff mean 30.2 of 46. Heterogeneity of measures included.	Low
Clarke et al (2015) [44]	There is some evidence that skills-based interventions presented in a module-based format can have a significant impact on promoting adolescent mental health and that cCBT has significant positive effects on adolescents' anxiety and depression symptoms; however, research is limited. Improvements of symptoms are maintained at 6 and 12 months.	Face-to-face and web-based support are associated with improved program completion and outcomes.	Not discussed	Limited (China)	On the basis of Quality Assessment Tool for Quantitative Studies, quality varied significantly from weak (12/20) to moderate or strong (7/20). Issues include a small number of studies, poor sampling, and heterogeneity across interventions.	Low

[a]AMSTAR: A Measurement Tool to Assess Systematic Reviews.

[b]RCT: randomized controlled trial.

[c]N/A: not applicable. This is a systematic review of apps and not studies, and therefore, quality assessment is not applicable.

[d]cCBT: computerized cognitive behavioral therapy.

[e]SMD: standardized mean difference.

[f]CBT: cognitive behavioral therapy.

[g]NNT: number needed to treat.

In conclusion, converging evidence across reviews suggests that digital health interventions have a small to medium effect when compared with nonactive controls (ie, waitlist or placebo). When compared with active controls, digital health interventions appear to be comparable, although findings varied by targeted set of symptoms, with evidence of effectiveness most apparent for anxiety and depression and to a lesser extent for stress. Inconclusive results across other symptom types were because of the limited number of trials conducted to date.

Effectiveness of Clinical Interventions

Most systematic reviews and meta-analyses have reported findings across studies that test the effectiveness of the implementation of computerized cognitive behavioral therapy (cCBT) [33,34,38-42,44]. Investigations of digital mental health interventions other than cCBT are rare, and thus, our analysis on the effectiveness of digital clinical interventions across studies focuses exclusively on cCBT.

According to 4 reviews, there is no significant difference in the effectiveness between cCBT delivered through a digital platform and standard face-to-face cognitive behavioral therapy (CBT) [38,39,41]. However, there is some evidence of benefits compared with nonactive controls [31,34,35,39,42].

Ye et al [41] found no statistical difference between internet-based CBT and face-to-face interventions, suggesting that the digital format may retain effectiveness. However, when compared with nonactive controls, cCBT was effective in reducing anxiety symptoms (SMD −0.52; 95% CI −0.90 to −0.14) but not in reducing depression (SMD 0.16; 95% CI 0.44-0.12) [41]. A meta-analysis by Podina et al [38] found that cCBT was as effective as standard CBT (Hedges *g*=0.295) and more effective than waitlist (Hedges *g*=1.410) in reducing

anxiety symptoms. Similarly, Vigerland et al [43] found a moderate effect on social anxiety disorder compared with waitlist controls (Hedges *g*=0.62; 95% CI 0.41-0.84). In 2 separate trials, older participants were found to gain greater clinical benefits compared with younger participants (slope=0.514) [24,38].

Ebert et al [39] found that the overall mean effect size of cCBT on symptoms of anxiety or depression was Hedges *g*=0.72 (95% CI 0.55-0.90) at posttest after controlling the baseline levels. This effect is similar to the effect of traditional CBT for anxiety (0.66) and higher than that of CBT for the treatment of depression in youth (0.35). When compared with a nonactive control, cCBT was effective in targeting anxiety (Hedges *g*=0.68; 95% CI 0.45-0.92; *P*<.001) and depression (Hedges *g*=0.76; 95% CI 0.41-0.12; *P*<.001).

With regard to studies with mixed comparison groups (active and nonactive), Harrer et al [34] found cCBT interventions more effective than others (eg, relationship skills training and emotional disclosure) for some conditions (depression: Hedges *g*=0.28; 95% CI 0.15-0.40 vs Hedges *g*=0.04; 95% CI −0.23 to 0.30; number needed to treat [NNT]=6.41 vs 4.4.5 and anxiety: Hedges *g*=0.36; 95% CI 0.23-0.50 vs Hedges *g*=−0.06; 95% CI −0.46 to 0.35; NNT: 5 vs 29.41). Similarly, Clarke et al [44] found that module-based cCBT showed significant positive effects in reducing depression and anxiety, thoughts of self-harm, and hopelessness and in improving sense of control.

Pennant et al [40] demonstrated greater effects when cCBT is targeted to young people assessed at risk of anxiety or depression, in comparison with the general population of young people. Among young people with elevated depression or anxiety symptom scores, cCBT had positive effects on anxiety (SMD 0.77; 95% CI 1.45-0.09; number of studies, n=6; number of participants=220) and depression (SMD 0.62; 95% CI 1.13-0.11; n=7; number of participants=279), whereas in the general population of young people, effect sizes were smaller (anxiety: SMD 0.15; 95% CI 0.26-0.03; number of participants=1273 and depression: SMD 0.15; 95% CI 0.26-0.03; number of participants=1280) [33]. Similar findings were also found in 2 other systematic reviews [34,44].

With regard to non-cCBT interventions, a small effect size of attention bias modification programs for anxiety and depression was observed (n=8; Hedges *g*=0.41; 95% CI 0.08-0.73; *P*<.01), whereas no benefit of cognitive bias modification programs or other interventions over either passive or active control groups (other therapeutically active conditions, attention or placebo training conditions, and waitlist) was observed [42].

Effectiveness of Digital Platforms

Only 4 systematic reviews have reported findings on digital platforms used to deliver digital mental health services. These included social networking sites [37], mental health apps [18,29], and therapeutic video games [28].

A systematic review by Ridout and Campbell [37] on social networking sites targeting mental health found no evidence of improvement in anxiety or psychosis symptoms in young people, whereas it found improvements in enhancing mental health knowledge and the number of depressive symptoms. Among the sites, the review suggested that the closed Facebook-like moderated online social therapy platforms as well as the YBMen project that used Facebook was effective, although there was no evidence of the effectiveness of other social networking platforms (the MindMax and Ching Story) included in the review [37]. In another systematic review, Grist et al [18] found no evidence to support the effectiveness of apps designed for adolescents with mental health conditions.

One reason for the lack of effectiveness across specific platforms may be attributable to a limited evidence base for many of the interventions available. For example, a review of 121 anxiety apps available in app stores (Google and Apple) by Bry et al [29] found that only a limited number of these apps were evidence based. Only one-sixth of the apps included educational information on the definition, symptoms, and treatment of anxiety. Half had at least one evidence-based treatment component, and one-fourth had more than one evidence-based treatment component, such as exposure therapy; thought challenging or cognitive restructuring; or self-monitoring of one's thoughts, emotions, and behaviors. The majority of those that lacked any evidence-based components were mostly distraction tools, such as games, coloring activities, or other audio or visual activities, and more than half included relaxation exercises, which are currently rarely considered therapeutic for anxiety [29]. Evidence on the effectiveness of therapeutic video games was limited and mixed, as confirmed by Barnes and Prescott [28].

Irrespective of their effectiveness or link with evidence-based approaches, young people generally perceive their engagement with these platforms to range from neutral to helpful. Overall, a systematic narrative review of Pretorius et al [36] reported that young people's perception of the helpfulness of web-based resources ranged across the studies—from 80% of participants in a study indicating that speaking on the web had helped, to 40% reporting in another study that web-based resources had helped a little, to 59% reporting in a third study that web-based resources did not make things better or worse.

Factors Associated With Effectiveness and Adherence

Several systematic reviews and meta-analyses demonstrated that digital mental health interventions with an in-person element (ie, therapist, parent, and peer) were more effective than those that were fully automatized or self-administered.

In another systematic review, Grist et al [42] found a significant effect of therapist support (Cochran Q=27.28; *P*<.001) and parental involvement (Cochran Q=24.43; *P*<.001). In their analysis, the involvement of a therapist yielded higher effect sizes (n=9; Hedges *g*=0.87; 95% CI 0.68-1.06; *P*<.001) than predominantly self-administered (Hedges *g*=0.81; 95% CI −0.68 to 2.31; *P*=.29) or purely self-administered interventions (Hedges *g*=0.24; 95% CI 0.10-0.38; *P*<.001). Similar findings were also reported by Hollis et al [24].

Garrido et al [35] reported higher pooled effect sizes of digital mental health interventions for depression with supervision than those without supervision (studies with no intervention controls: Cohen *d*=0.52; 95% CI 0.23-0.80 and studies with active controls: Cohen *d*=0.49; 95% CI −0.11 to 1.01). In a systematic

review and meta-analysis, Valimaki et al [33] found that web-based interventions with a human element, including face-to-face guidance, monitoring of engagement, or follow-up telephone calls by teachers and health professionals, were more effective than those without a human element.

Grist et al [42] demonstrated a significant difference in effect sizes (Cochran Q=9.37; *P*=.002) between trials with ongoing psychological or pharmacological treatment (Hedges *g*=0.90, 95% CI 0.68-1.11; *P*<.001) and trials without ongoing treatment (Hedges *g*=0.42, 95% CI 0.20-0.63).

In contrast, Harrer et al [34] did not find supervision significantly affecting intervention efficacy; however, this may be because of the multiplicity of the types of interventions included in the review or the older target population (university students). In addition, Ebert et al [39] found no association between parental involvement and better treatment outcomes of cCBT for anxiety or depression in youth (without parental involvement: Hedges *g*=0.83, 95% CI 0.53-1.13; *P*<.001; NNT=2.26 and with parental involvement: Hedges *g*=0.64, 95% CI 0.40-0.88; *P*<.001; NNT=2.86).

An in-person element was also associated with adherence and lower dropout rates. Clarke et al [44] suggested that face-to-face or web-based support in web-based interventions was associated with better completion and outcomes. Similarly, Hollis et al [24] reported that human involvement is positively associated with adherence; however, they note that the evidence is scant.

Human contact in digital mental health interventions was also considered useful and valuable by adolescents and young people themselves, in particular, contact with professionals as well as peers with similar experiences and mental health issues [35,37]. Pretorius et al [36] found that young people valued web-based services run by mental health professionals and the opportunity to connect to peers, with 84% of participants reporting that human contact within a web-based mental health resource is important. In addition, in a systematic review by Ridout and Campbell [37], the involvement of professionals and peers in social networking sites was valued by site users.

Hollis et al [24] reported that adolescents and young people prefer face-to-face mental health interventions over digital interventions. In the Australian sample, two-thirds (59%) of young people strongly preferred face-to-face treatment, with only 16% preferring on the web, and in the United Kingdom, half were not interested in cCBT, with preference for face-to-face treatment.

There was some indication that interventions implemented in the school setting were associated with improvements in adolescent mental health knowledge, support seeking, and well-being [44]. School- and web-based interventions were also associated with greater adherence [35,44], and interventions that adolescents and young people completed in their own time were associated with low completion rates and adherence [35].

Design Elements

Acceptability of interventions was reported to be good [30,37,40,44]. Privacy, safety, and discretion were found to be valuable for adolescents and young people [24,30,36]. Related to the stigma associated with mental health issues, adolescents and young people also valued anonymity [24,36]. In this regard, data security, including password protection, control over privacy settings [30], and site moderation by professionals [36,37], were identified as factors influencing the acceptability of digital mental health interventions.

Other characteristics valued by adolescents and young people included the credibility of design, visual appearance, and information and resources provided [30,35,36]. The tools and content should be engaging and interactive [30,35]; should provide concise, interesting, and trustworthy information [24,30,35,36]; should be esthetically attractive [35,36]; should provide reminders to use [30]; should allow for personalization [30]; should have relatable situations, characters, or avatars [35]; and should reflect local and cultural differences and needs, particularly in terms of minority groups and migrants for social integration. Garrido et al [35] reported that technical glitches were a barrier to complete interventions.

Flexibility, self-reliance, and control were also cited in the reviews as influencing acceptability [24,36]. Adolescents and young people valued in digital mental health interventions the ability to complete interventions on their own terms and pace [24]. According to Pretorius et al [36], 24-hour availability is an important factor, as help-seeking takes place mostly after 11 PM.

Sustainability, Completion, and Adherence

Most studies included in this review reported only short-term effects on adolescents' mental health. Evidence of long-term effects is limited [24,33,38-41,44]. Only one meta-analysis by Valimaki et al [33] with a focus on depression, anxiety, and stress examined the long-term effects of digital health interventions. The study found a statistically significant improvement at the end of the intervention on depressive symptoms (*P*=.02; median 1.68, 95% CI 3.11-0.25) and after 6 months (*P*=.01; median 1.78, 95% CI 3.20-0.37). The study also found evidence of long-term improvement at 6 months in anxiety symptoms (*P*<.001; median 1.47, 95% CI 2.36-0.59) and moods and feelings (*P*=.04; median 5.55, 95% CI 10.88-0.22), but there was no difference in stress scores.

In terms of cCBT, in line with the standard CBT, effects were higher for interventions of moderate length (1-2 months), for example, on depression at 4-8 weeks (Hedges *g*=0.31, 95% CI 0.13-0.49; NNT=5.75) compared with shorter (Hedges *g*=0.09, 95% CI −0.02 to 0.21; NNT=20) or longer (Hedges *g*=0.13, 95% CI −0.43 to 0.69; NNT=13.51) programs (*P*=.03), according to Harrer et al [34]. Although follow-up assessments were rarely reported in studies, Clarke et al [44] also found that improvements after cCBT were maintained at 6 and 12 months.

In addition to limited evidence of the long-term effects of digital mental health interventions, Hollis et al [24] found limited evidence of a dose-response (ie, how much of the intervention is needed to produce beneficial outcomes).

Overall, dropout was found to be high in the systematic reviews and meta-analyses of studies on digital mental health interventions. Completion rates ranged greatly from 10% to 94% in a study by Valimaki et al [33] and from 65% to 83%

among app users in a study by Grist et al [30], and completion rates were approximately half on average in a study by Clarke et al [44]. However, data on dropout and adherence were generally considered weak in the original review samples, with only a limited number of studies reporting data on adherence [24,30,33,35,36,44].

Gender was considered as a predictor of adherence. According to Garrido et al [35], females were more likely to complete the intervention than males [35]. In addition, mental health status was associated with completion, and higher completion was predicted for adolescents and young people with higher depression scores at the baseline [35,44], a longer history of mood disorders, or low anxiety scores at pretest [36]. Furthermore, according to Pretorius et al [36], high levels of psychological distress were associated with help-seeking on the web.

Cost-Effectiveness

Data on cost-effectiveness were not reported in any of the systematic reviews and meta-analyses in our sample, and there was no indication of research and development costs. A total of 5 systematic reviews noted that despite being widely considered low cost, for example, because of reduced time and personnel expenses [43], there is still a lack of data on the cost-effectiveness and economic benefits of digital mental health interventions [24,29,32,33,42].

Generalizability of Findings

None of the studies reported on the socioeconomic background or other characteristics of the target populations. Most studies were conducted in high-income countries across Europe (n=71) and in the United States (n=21), Australia (n=21), Canada (n=13), and New Zealand (n=9). In terms of low- and middle-income economies, interventions were reported only from 4 countries, with most studies conducted in China (n=9), including Hong Kong, and, to a lesser extent, in Chile (n=2), Egypt (n=1), and Thailand (n=1). Given the homogeneity of the country contexts and lack of analysis of the characteristics of the target population, the generalizability of the findings is limited beyond adolescents and young people in high-income country settings.

Discussion

Principal Findings

We explored 18 reviews and meta-analyses on the effectiveness of digital mental health interventions for adolescents and young people. On the basis of this systematic overview, we found evidence on the effectiveness of cCBT on anxiety and depression, whereas the effectiveness of other digital mental health interventions, including therapeutic video games, mobile apps, or social networking sites, remains inconclusive. The effects vary based on a targeted set of symptoms, with evidence of effectiveness found on anxiety; depression; and, to a lesser extent, stress, and based on age, with older participants gaining greater benefits compared with younger adolescent participants.

Digital interventions that deploy evidence-based treatment such as cCBT are generally comparable with face-to-face care. Importantly, in-person elements (eg, professional, peer, or parent engagement) were found to strengthen the effectiveness of digital interventions. In addition, digital interventions improved outcomes relative to waitlist controls, suggesting that they may have additional benefits for supporting adolescents and young people in cases where access to care is limited or wait times to access are long.

Furthermore, although young people report a range of neutral to positive attitudes about the helpfulness of digital platforms for mental health support, few studies have tracked the long-term outcomes of digital mental health interventions. Although acceptability is considered good, dropout is common, and adherence is relatively weak if not boosted by in-person elements. Very little is known about cost-effectiveness, with no systematic reviews or meta-analyses reporting on cost-effectiveness. Finally, given that the vast majority of interventions are implemented in high-income countries, very little is known about the generalizability of the findings to low- and middle-income countries and to a range of adolescents and young people with different socioeconomic, cultural, racial, or other backgrounds.

Despite some converging evidence across meta-analyses and reviews, research in this area appears to have consistently low quality and rigor as per assessment using the AMSTAR 2 criteria. The primary constraints for this were that the articles analyzed reported many limitations in their samples. These included a small number of studies meeting the inclusion criteria [28,30,31,38,39,41,44], weak quality of studies [32,34,40,41,43,44], and the heterogeneity across the interventions in terms of content and delivery [24,31,34,36,39-41,44]. Furthermore, study participants were often recruited by self-selection [30,37,44], sample sizes were small [24,30,32], and blinding was limited [24,30,35]. Notably, one systematic review by Grist et al [30] also pointed out that almost all studies were either undertaken or supported by the program developer, which may greatly affect the study design and interpretation of the findings.

Comparison With Previous Work

With the growing application of digital technologies in public health, digital health interventions are perceived to increase access to health services and information, self-care, and empowerment and reduce the cost and burden on health systems [45]. In this context, as *digital natives*, adolescents and young people are considered as early adopters of technology [46], with the potential to benefit from digital health technologies, including for mental health.

Although there is an increasing body of research on the effectiveness of digital mental health technologies targeting adolescents and young people, most focus on evaluating cCBT. In line with our findings, cCBT for addressing anxiety and depression in adolescents and young people has been found to be effective, including in school-based prevention and early identification studies and in family-based studies [47]. The effectiveness of cCBT in the adult population has also been established [48,49]. Given that face-to-face CBT is widely used as a treatment for depressive symptoms and disorders in this age group [50], with evidence of its effectiveness found in a

number of systematic reviews [51-53], it is plausible that it also works in a standardized digital format.

Beyond the cCBT, evidence on the effectiveness of other digital mental health interventions, including therapeutic video games, mobile apps, and social networking sites, was extremely limited. Although these may have the potential to engage adolescents and young people and thus support traditional face-to-face treatment [54] and although social network sites, including gaming elements, are found to be promising in promoting changes in health-related behaviors [55,56], the quality of content and expected outcomes vary [57].

Similar to our findings, studies have reported low adherence and high dropout rates in adolescents and young people using digital mental health interventions [58-60], although there are also contrasting data with high levels of acceptability and usability [47], including from low- and middle-income countries [61-63]. However, the contrasting data are mainly reported in feasibility studies, based on adolescents without mental health conditions, and thus, the data may not be applicable for adolescents and young people with mental health issues.

Furthermore, to some extent, the cost-effectiveness of digital health interventions has been studied in the general population and other areas of health, including the management of cardiovascular diseases [64] and insomnia [65]. However, there is a lack of assessment of cost-effectiveness in digital mental health interventions overall and for adolescents and young people in particular. This may be because of methodological limitations related to a number of studies, including heterogeneity of interventions and outcomes that hinder the overall assessment of effectiveness.

Finally, despite an increasing share of young population and users of digital technology in low- and middle-income countries, very little research has been conducted in these settings [61,62]. In line with our findings, the generalizability to low- and middle-income countries [47,66] as well as adolescents and young people with different backgrounds [47] is noted by previous research. However, good-quality research on cost-effectiveness and generalizability is critical when scaling up these interventions in settings with already limited resources for health care, including mental health services.

Limitations

Although this overview of meta-analyses and systematic reviews provides a broad assessment of the results and quality of digital mental health intervention research focused on adolescents and young people, several limitations are evident. In this overview, we have provided a higher-level synthesis of previous systematic reviews in this area, covering a range of digital health interventions and expected health outcomes. Although this is a critical step in assessing the value of digital interventions overall, it introduces some challenges for interpretation (eg, variation in study settings, methods, and comparators, with inconsistencies in reporting within and across the reviews, including the level of description of primary studies and the findings). However, these inconsistencies highlight an important need for more systematic approaches to testing and reporting on effectiveness across studies. Inclusion criteria for some of the studies reviewed here may have resulted in overlap of primary studies between the reviews. In addition, as the field of digital interventions is fast-moving, many of the interventions tested may now be outdated or defunct. However, cross-study heterogeneity is why this review is needed to identify converging effects that emerge, despite variation in specific tests across studies and reviews.

Finally, we included only published peer-reviewed systematic reviews and meta-analyses in the English language. Inclusion of randomized control trials and other original research, including in other languages, may have yielded more studies focused on low- and middle-income countries.

Despite these limitations, the present overview provides a broad picture of the converging evidence supporting the promise of digital mental health interventions in adolescents and young people and highlights a critical need for the field to increase the number of high-quality effectiveness trials to ensure that the interest and enthusiasm in these approaches do not outpace their results.

Conclusions

This overview of meta-analyses and systematic reviews suggests that digital mental health interventions for adolescents and young people have modest positive effects, especially when relying on evidence-based treatment content or in-person elements that boost engagement. Their potential for settings with limited resources for health and cost savings compared with traditional treatment remains understudied. Therefore, when developing, investing in, and delivering digital mental health programs for adolescents and young people, we need to better consider what types of services are meaningful to be provided through a digital platform (ie, cCBT that deploys the same techniques as face-to-face therapy and is typically delivered by a professional), for what outcomes (eg, self-reported vs diagnosed and mild vs severe symptoms), what type of services adolescents and young people themselves prefer (standard vs digital), and to what extent these are cost-saving and clinically effective across a variety of settings with different resources (ie, in high- vs low-resource settings).

Acknowledgments

This work was supported by the Fondation Botnar, Basel, Switzerland. The authors thank Silvia Minozzi and David A Ross for their advice on the structure and content of this review.

Authors' Contributions

SL contributed to conceptualizing and designing the study; conducted literature search, quality assessment, analysis, and data interpretation; drafted the initial manuscript; and reviewed and revised the manuscript. KTF contributed to advancing the ideas, aided with analytic interpretation, and reviewed and revised the manuscript. JM contributed to the literature search, quality assessment, data extraction, analysis, and review of the manuscript. BW contributed to conceptualizing and designing the study. NS contributed to conceptualizing and designing the study, conducted analysis and data interpretation, drafted the initial manuscript, and reviewed and revised the manuscript. SL and NS are joint first authors of the review.

Conflicts of Interest

None declared.

References

1. WHO. Adolescent Mental Health. Fact Sheet. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [accessed 2020-07-28]
2. Patton GC, Sawyer SM, Santelli JS, Ross DA, Afifi R, Allen NB, et al. Our future: a Lancet commission on adolescent health and wellbeing. Lancet 2016 Jun 11;387(10036):2423-2478. [doi: 10.1016/S0140-6736(16)00579-1] [Medline: 27174304]
3. WHO. Adolescent mental health. Published October 23. 2019. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [accessed 2020-06-08]
4. UNICEF. Adolescent mental health. Published August. 2019. URL: https://data.unicef.org/topic/child-health/mental-health/ [accessed 2020-06-08]
5. WHO. Suicides by age and income level. Published. 2016. URL: https://www.who.int/mental_health/prevention/suicide/suicideprevent/en/ [accessed 2020-06-08]
6. WHO. Causes of Death among Adolescents. Global Health Estimates (GHE). WHO. 2016. URL: https://www.who.int/healthinfo/global_burden_disease/estimates/en/ [accessed 2020-07-06] [WebCite Cache ID https://www.who.int/healthinfo/global_burden_disease/estimates/en/]
7. UNICEF. Growing up in a Connected World. UNICEF Office of Research ? Innocenti. 2019. URL: https://www.unicef-irc.org/publications/1060-growing-up-in-a-connected-world.html [accessed 2020-06-09]
8. UNICEF. Children in a Digital World. In: The State of the World's Children. New York: UNICEF; 2017.
9. International TU, World TDR. Mobile cellular subscriptions; individuals using the Internet. Published. 2018. URL: https://data.worldbank.org/indicator/IT.CEL.SETS.P2 [accessed 2020-07-28]
10. Israni ST, Matheny ME, Matlow R, Whicher D. Equity, Inclusivity, and Innovative Digital Technologies to Improve Adolescent and Young Adult Health. J Adolesc Health 2020 Aug;67(2S):S4-S6 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.05.014] [Medline: 32718514]
11. Webb L, Clary LK, Johnson RM, Mendelson T. Electronic and School Bullying Victimization by Race/Ethnicity and Sexual Minority Status in a Nationally Representative Adolescent Sample. J Adolesc Health 2020 Jul 10:378-384. [doi: 10.1016/j.jadohealth.2020.05.042] [Medline: 32654837]
12. Moreno MA, Kerr BR, Jenkins M, Lam E, Malik FS. Perspectives on Smartphone Ownership and Use by Early Adolescents. J Adolesc Health 2019 Apr;64(4):437-442. [doi: 10.1016/j.jadohealth.2018.08.017] [Medline: 30389204]
13. Giovanelli A, Ozer EM, Dahl RE. Leveraging Technology to Improve Health in Adolescence: A Developmental Science Perspective. J Adolesc Health 2020 Aug;67(2S):S7-S13 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.02.020] [Medline: 32718517]
14. Aschbrenner KA, Naslund JA, Tomlinson EF, Kinney A, Pratt SI, Brunette MF. Adolescents' Use of Digital Technologies and Preferences for Mobile Health Coaching in Public Mental Health Settings. Front Public Health 2019;7:178 [FREE Full text] [doi: 10.3389/fpubh.2019.00178] [Medline: 31312629]
15. Naslund JA, Aschbrenner KA, McHugo GJ, Unützer J, Marsch LA, Bartels SJ. Exploring opportunities to support mental health care using social media: A survey of social media users with mental illness. Early Interv Psychiatry 2017 Oct 20;13(3):405-413. [doi: 10.1111/eip.12496] [Medline: 29052947]
16. Shortliffe EH. Digital Medicine and Biomedical Informatics: What's in a Name? Methods Inf Med 2016 Oct 17;55(5):389-391. [doi: 10.3414/ME15-22-0005] [Medline: 27524308]
17. Kaushik A, Kostaki E, Kyriakopoulos M. The stigma of mental illness in children and adolescents: A systematic review. Psychiatry Res 2016 Sep 30;243:469-494. [doi: 10.1016/j.psychres.2016.04.042] [Medline: 27517643]
18. Grist R, Porter J, Stallard P. Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
19. WHO. Health for the World's Adolescents: A Second Chance in the Second Decade. Geneva: WHO; 2014:A.
20. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. Lancet Child Adolesc Health 2018 Mar;2(3):223-228. [doi: 10.1016/S2352-4642(18)30022-1] [Medline: 30169257]

21. WHO. Mental Disorders. WHO. 2019. URL: https://www.who.int/en/news-room/fact-sheets/detail/mental-disorders [accessed 2020-10-02]
22. WHO. Universal Health Coverage for Mental Health. Geneva: WHO; 2019.
23. Lenhard F, Andersson E, Mataix-Cols D, Rück C, Vigerland S, Högström J, et al. Therapist-Guided, Internet-Delivered Cognitive-Behavioral Therapy for Adolescents With Obsessive-Compulsive Disorder: A Randomized Controlled Trial. J Am Acad Child Adolesc Psychiatry 2017 Jan;56(1):10-19.e2 [FREE Full text] [doi: 10.1016/j.jaac.2016.09.515] [Medline: 27993223]
24. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual Research Review: Digital health interventions for children and young people with mental health problems: a systematic and meta-review. J Child Psychol Psychiatry 2017;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
25. Merriam-Webster CD. Cognitive Behavioural Therapy, CBT. 2020. URL: https://www.merriam-webster.com/dictionary/cognitive%20behavioral%20therapy [accessed 2020-09-02]
26. Karlsson P, Bergmark A. Compared with what? An analysis of control-group types in Cochrane and Campbell reviews of psychosocial treatment efficacy with substance use disorders. Addiction 2015 Jan 15;110(3):420-428. [doi: 10.1111/add.12799]
27. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017 Dec 21;358:j4008 [FREE Full text] [doi: 10.1136/bmj.j4008] [Medline: 28935701]
28. Barnes S, Prescott J. Empirical Evidence for the Outcomes of Therapeutic Video Games for Adolescents With Anxiety Disorders: Systematic Review. JMIR Serious Games 2018 Feb 28;6(1):e3 [FREE Full text] [doi: 10.2196/games.9530] [Medline: 29490893]
29. Bry LJ, Chou T, Miguel E, Comer JS. Consumer Smartphone Apps Marketed for Child and Adolescent Anxiety: A Systematic Review and Content Analysis. Behav Ther 2018 Dec;49(2):249-261. [doi: 10.1016/j.beth.2017.07.008] [Medline: 29530263]
30. Grist R, Porter J, Stallard P. Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
31. Davies EB, Morriss R, Glazebrook C. Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: a systematic review and meta-analysis. J Med Internet Res 2014;16(5):e130 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]
32. Farrer L, Gulliver A, Chan JKY, Batterham PJ, Reynolds J, Calear A, et al. Technology-based interventions for mental health in tertiary students: systematic review. J Med Internet Res 2013;15(5):e101 [FREE Full text] [doi: 10.2196/jmir.2639] [Medline: 23711740]
33. Välimäki M, Anttila K, Anttila M, Lahti M. Web-Based Interventions Supporting Adolescents and Young People With Depressive Symptoms: Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2017 Dec 08;5(12):e180 [FREE Full text] [doi: 10.2196/mhealth.8624] [Medline: 29222079]
34. Harrer M, Adam SH, Baumeister H, Cuijpers P, Karyotaki E, Auerbach RP, et al. Internet interventions for mental health in university students: A systematic review and meta-analysis. Int J Methods Psychiatr Res 2019 Jun;28(2):e1759. [doi: 10.1002/mpr.1759] [Medline: 30585363]
35. Garrido S, Millington C, Cheers D, Boydell K, Schubert E, Meade T, et al. What Works and What Doesn't Work? A Systematic Review of Digital Mental Health Interventions for Depression and Anxiety in Young People. Front Psychiatry 2019;10:759 [FREE Full text] [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
36. Pretorius C, Chambers D, Coyle D. Young People's Online Help-Seeking and Mental Health Difficulties: Systematic Narrative Review. J Med Internet Res 2019 Nov 19;21(11):e13873 [FREE Full text] [doi: 10.2196/13873] [Medline: 31742562]
37. Ridout B, Campbell A. The Use of Social Networking Sites in Mental Health Interventions for Young People: Systematic Review. J Med Internet Res 2018 Dec 18;20(12):e12244 [FREE Full text] [doi: 10.2196/12244] [Medline: 30563811]
38. Podina IR, Mogoase C, David D, Szentagotai A, Dobrean A. A Meta-Analysis on the Efficacy of Technology Mediated CBT for Anxious Children and Adolescents. J Rat-Emo Cognitive-Behav Ther 2015 Nov 17;34(1):31-50. [doi: 10.1007/s10942-015-0228-5]
39. Ebert DD, Zarski A, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. PLoS One 2015;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]
40. Pennant ME, Loucas CE, Whittington C, Creswell C, Fonagy P, Fuggle P, et al. Computerised therapies for anxiety and depression in children and young people: a systematic review and meta-analysis. Behav Res Ther 2015 Apr;67:1-18. [doi: 10.1016/j.brat.2015.01.009] [Medline: 25727678]
41. Ye X, Bapuji SB, Winters SE, Struthers A, Raynard M, Metge C, et al. Effectiveness of internet-based interventions for children, youth, and young adults with anxiety and/or depression: a systematic review and meta-analysis. BMC Health Serv Res 2014 Jul 18;14:313 [FREE Full text] [doi: 10.1186/1472-6963-14-313] [Medline: 25037951]

42. Grist R, Croker A, Denne M, Stallard P. Technology Delivered Interventions for Depression and Anxiety in Children and Adolescents: A Systematic Review and Meta-analysis. Clin Child Fam Psychol Rev 2019 Jun;22(2):147-171 [FREE Full text] [doi: 10.1007/s10567-018-0271-8] [Medline: 30229343]
43. Vigerland S, Lenhard F, Bonnert M, Lalouni M, Hedman E, Ahlen J, et al. Internet-delivered cognitive behavior therapy for children and adolescents: A systematic review and meta-analysis. Clin Psychol Rev 2016 Dec;50:1-10 [FREE Full text] [doi: 10.1016/j.cpr.2016.09.005] [Medline: 27668988]
44. Clarke AM, Kuosmanen T, Barry MM. A systematic review of online youth mental health promotion and prevention interventions. J Youth Adolesc 2015 Jan;44(1):90-113. [doi: 10.1007/s10964-014-0165-0] [Medline: 25115460]
45. WHO. WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. Geneva: WHO; 2019.
46. Giovanelli A, Ozer EM, Dahl RE. Leveraging Technology to Improve Health in Adolescence: A Developmental Science Perspective. J Adolesc Health 2020 Aug;67(2S):S7-S13 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.02.020] [Medline: 32718517]
47. Das JK, Salam RA, Lassi ZS, Khan MN, Mahmood W, Patel V, et al. Interventions for Adolescent Mental Health: An Overview of Systematic Reviews. Journal of Adolescent Health 2016 Oct;59(4):S49-S60. [doi: 10.1016/j.jadohealth.2016.06.020]
48. Adelman CB, Panza KE, Bartley CA, Bontempo A, Bloch MH. A meta-analysis of computerized cognitive-behavioral therapy for the treatment of DSM-5 anxiety disorders. J Clin Psychiatry 2014 Jul;75(7):e695-e704. [doi: 10.4088/JCP.13r08894] [Medline: 25093485]
49. Andrews G, Basu A, Cuijpers P, Craske M, McEvoy P, English C, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. Journal of Anxiety Disorders 2018 Apr;55:70-78. [doi: 10.1016/j.janxdis.2018.01.001]
50. Oud M, de Winter L, Vermeulen-Smit E, Bodden D, Nauta M, Stone L, et al. Effectiveness of CBT for children and adolescents with depression: A systematic review and meta-regression analysis. Eur. psychiatr 2019 Jan 16;57:33-45. [doi: 10.1016/j.eurpsy.2018.12.008]
51. Sigurvinsdóttir AL, Jensínudóttir KB, Baldvinsdóttir KD, Smárason O, Skarphedinsson G. Effectiveness of cognitive behavioral therapy (CBT) for child and adolescent anxiety disorders across different CBT modalities and comparisons: a systematic review and meta-analysis. Nordic Journal of Psychiatry 2019 Nov 18;74(3):168-180. [doi: 10.1080/08039488.2019.1686653]
52. Viswanathan M, Kennedy S, McKeeman J, Christian R. Treatment of Depression in Children and Adolescents. A Systematic Review. AHRQ Comparative Effectiveness Reviews. Report No.EHC005-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2020.
53. Treatment of Anxiety Disorders. A Systematic Review. SBU Yellow Report No. 171/1+2. Stockholm: Swedish Council on Health Technology Assessment; 2005.
54. Al Husni AKM, Delvenne V. Use of Video Games in a Child And Adolescent Psychiatric Unit. Psychiatr Danub. 167AD;32(Suppl 1) 2020:167-167. [Medline: 32890384]
55. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256. [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
56. Chau MM, Burgermaster M, Mamykina L. The use of social media in nutrition interventions for adolescents and young adults—A systematic review. International Journal of Medical Informatics 2018 Dec;120:77-91. [doi: 10.1016/j.ijmedinf.2018.10.001]
57. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
58. Achilles MR, Anderson M, Li SH, Subotic-Kerry M, Parker B, O'Dea B. Adherence to e-mental health among youth: Considerations for intervention development and research design. DIGITAL HEALTH 2020 May 21;6:205520762092606. [doi: 10.1177/2055207620926064]
59. Liverpool S, Mota CP, Sales CMD, 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. [doi: 10.2196/16317]
60. Lillevoll KR, Vangberg HCB, Griffiths KM, Waterloo K, Eisemann MR. Uptake and adherence of a self-directed internet-based mental health intervention with tailored e-mail reminders in senior high schools in Norway. BMC Psychiatry 2014;14(1):14. [doi: 10.1186/1471-244X-14-14]
61. Sobowale K, Nguyen M, Weiss B, Van TTH, Trung LT. Acceptability of internet interventions for youth mental health in Vietnam. Glob. Ment. Health 2016 Jul 14;14(3):e22. [doi: 10.1017/gmh.2016.18]
62. Kanuri N, Arora P, Talluru S, Colaco B, Dutta R, Rawat A, et al. Examining the initial usability, acceptability and feasibility of a digital mental health intervention for college students in India. Int J Psychol 2019 Dec 22;55(4):657-673. [doi: 10.1002/ijop.12640]

63. Maloney CA, Abel WD, McLeod HJ. Jamaican adolescents' receptiveness to digital mental health services: A cross-sectional survey from rural and urban communities. Internet Interventions 2020 Sep;21(Sept 21):100325. [doi: 10.1016/j.invent.2020.100325]
64. Jiang X, Ming W, You JH. The Cost-Effectiveness of Digital Health Interventions on the Management of Cardiovascular Diseases: Systematic Review. J Med Internet Res 2019 Jun 17;21(6):e13166. [doi: 10.2196/13166]
65. Darden M, Espie C, Carl J, Henry A. Cost-effectiveness of digital CBT (Sleepio) for insomnia: A Markov simulation model in the United States. Sleep 2020:zsaa223. [doi: 10.1093/sleep/zsaa223]
66. Abaza H, Marschollek M. mHealth Application Areas and Technology Combinations. A Comparison of Literature from High and Low/Middle Income Countries. Methods Inf Med 2017 Aug 08;56(7):e105-e122. [doi: 10.3414/ME17-05-0003] [Medline: 28925418]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
AMSTAR 2: A Measurement Tool to Assess Systematic Reviews
CBT: cognitive behavioral therapy
cCBT: computerized cognitive behavioral therapy
NNT: number needed to treat
SMD: standardized mean difference

Edited by J Torous; submitted 18.11.20; peer-reviewed by M Brunette, M Bonnert; comments to author 22.12.20; revised version received 15.02.21; accepted 10.03.21; published 29.04.21.

Please cite as:
Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N
Evidence on Digital Mental Health Interventions for Adolescents and Young People: Systematic Overview
JMIR Ment Health 2021;8(4):e25847
URL: https://mental.jmir.org/2021/4/e25847
doi:10.2196/25847
PMID:33913817

Original Paper

Preparing Patients and Clinicians for Open Notes in Mental Health: Qualitative Inquiry of International Experts

Charlotte Blease[1], PhD; John Torous[2], MD, MBI; Anna Kharko[3], BSc; Catherine M DesRoches[1,4], DrPH; Kendall Harcourt[1], MPA; Stephen O'Neill[1], LICSW, BCD, JD; Liz Salmi[1], AS; Deborah Wachenheim[1], MPP; Maria Hägglund[5], PhD

[1]Division of General Medicine, Beth Israel Deaconess Medical Center, Boston, MA, United States

[2]Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, United States

[3]School of Psychology, University of Plymouth, Plymouth, United Kingdom

[4]Department of Psychiatry, Harvard Medical School, Boston, MA, United States

[5]Department of Women's and Children's Health, University of Uppsala, Uppsala, Sweden

Corresponding Author:
Charlotte Blease, PhD
Division of General Medicine
Beth Israel Deaconess Medical Center
133 Brookline Avenue
Boston, MA, 02115
United States
Phone: 1 16173201281
Email: charlotteblease@gmail.com

Abstract

Background: In a growing number of countries worldwide, clinicians are sharing mental health notes, including psychiatry and psychotherapy notes, with patients.

Objective: The aim of this study is to solicit the views of experts on provider policies and patient and clinician training or guidance in relation to open notes in mental health care.

Methods: In August 2020, we conducted a web-based survey of international experts on the practice of sharing mental health notes. Experts were identified as informaticians, clinicians, chief medical information officers, patients, and patient advocates who have extensive research knowledge about or experience of providing access to or having access to mental health notes. This study undertook a qualitative descriptive analysis of experts' written responses and opinions (*comments*) to open-ended questions on training clinicians, patient guidance, and suggested policy regulations.

Results: A total of 70 of 92 (76%) experts from 6 countries responded. We identified four major themes related to opening mental health notes to patients: the need for clarity about provider policies on exemptions, providing patients with basic information about open notes, clinician training in writing mental health notes, and managing patient-clinician disagreement about mental health notes.

Conclusions: This study provides timely information on policy and training recommendations derived from a wide range of international experts on how to prepare clinicians and patients for open notes in mental health. The results of this study point to the need for further refinement of exemption policies in relation to sharing mental health notes, guidance for patients, and curricular changes for students and clinicians as well as improvements aimed at enhancing patient and clinician-friendly portal design.

(JMIR Ment Health 2021;8(4):e27397) doi:10.2196/27397

KEYWORDS

open notes; electronic health records; attitudes; survey; mental health; psychiatry; psychotherapy; qualitative research; mobile phone

Introduction

Background

A growing number of health organizations worldwide now offer patients web-based access to their clinical notes (*open notes*) [1]. Using secure internet portals, patients can log in and, at their own convenience, rapidly access and read their clinical documentation. Access is not just to lists of medications, laboratory results, referrals, or visit summaries but also to the very words written by clinicians. Emerging from the participatory design movement in Scandinavia [2]—often called *the Scandinavian Approach*—which strives for *democracy and democratization* in digital design [3], most patients in the Nordic countries are already offered open notes. Different cultural and policy considerations have advanced open notes in the United States; starting April 5, 2021 (postponed from November 2, 2020, owing to COVID-19), new federal rules mandate that, with few exceptions, all health providers must offer patients access to their web-based clinical notes [4,5]. The driving force behind these rules is the 21st-century Cures Act, which was enacted with the goal of accelerating medical product development and bringing innovations to patients quickly and efficiently [6].

However, sharing mental health notes, including psychiatry and psychotherapy notes, remains controversial, and some clinicians are uncertain about when it is appropriate to hide notes from patients. In Sweden, where the majority of patients have access to open notes (via *Journalen* the country's eHealth portal), mental health notes are shared in psychiatric centers in 11 out of 21 regions of the country. In Norway, where all patients can access their clinical notes, a survey of psychiatry clinicians working in hospitals found that 8% kept a *shadow record* to prevent patients from reading all of their notes [7]. In the United States, psychotherapy notes are exempt from the new federal rules, and *information blocking* is permitted, if doing so "...will substantially reduce the risk of harm" to a patient or to another person [4]. Licensed health professionals can decide what constitutes a substantial risk "...in the context of a current or prior clinician-patient relationship" [4]. These rules leave considerable room for interpretation, and it is unclear how clinicians' discretion will be monitored or evaluated.

Beyond policy and auditing considerations, surveys show that many mental health clinicians worry that patients will become anxious or confused after reading their notes [8,9]. In a study conducted at the US Veterans Health Administration, 63% (n=127) of clinician respondents described being less detailed in their documentation as a result of patient access and 49% (n=98) reported that they would be *pleased* if the practice was discontinued [9]. In Sweden, 62% (n=39) of clinical psychologists reported being less candid in their notes after the implementation of patient access to Journalen [8]. In lieu of adequate clinician training on writing mental health notes, ad hoc strategies aimed at minimizing patient harm, confusion, or disagreements with clinicians may undermine best practice.

Some surveys of patients' experiences with reading their mental health notes are promising. For example, in a recent comparison of primary care patients with and without a mental health diagnosis, Klein et al [10] reported no differences in patient experiences with open notes: 92% (336) of patients with a mental health diagnosis compared with 91% (1789) of patients without a mental health diagnosis reported feeling more in control of their health care. A pilot study at an outpatient psychiatric clinic in Boston found that the majority of patients reported a better understanding of their mental health condition and better remembering their care plan [11]. However, not all patients report benefits from reading their mental health notes, and some studies suggest that patient trust may be enhanced or strained by access [12,13] Generally, research on opening notes with mental health patients is limited, and there is little discussion in the literature about how to provide patient guidance about the benefits and risks of accessing their clinical notes [14].

Objectives

There is now extensive research on sharing outpatient visit notes with patients seeking medical care [15,16]. Although some surveys have examined mental health clinicians' attitudes about open notes [9,11], only a few have been conducted among clinicians with experience of sharing their notes with patients [8,17]. Although in many countries, such as the United States, the majority of mental health care is provided in primary care, so far, few surveys have analyzed the experiences of open notes among mental health patients in that setting [10]. In addition, few surveys have solicited the experiences and opinions of patients who have read outpatient or inpatient psychiatry or psychotherapy notes [11,13,18,19]. Relatedly, we are not aware of any studies that have set out to explore the experiences with open notes of patients living with serious mental illnesses such as psychotic disorders, major depression, and bipolar disorders. Finally, only a limited number of investigations have examined how to prepare clinicians and patients for opening notes in the context of mental health care [14,20,21].

As previous publications have emphasized, open notes in mental health care do raise new practice dilemmas [18,22,23]. Clinicians must balance the duty to respect patient autonomy and transparent information disclosures while preventing the potential for patient harm from reading notes that may be upsetting or confusing [22,24]. Considering the pressing need for greater clarity about best practice in this domain, our goal was to initiate expert-led discussion on policy recommendations, including on how to better train clinicians and guide patients, for this practice innovation.

Methods

Background

We used a structured web-based survey to explore the consensus views of international experts. The qualitative web survey was embedded in a modified Delphi methodology structured around 3 rounds of surveys. The Delphi technique is an established methodology for exploring the consensus views of experts. It is especially well suited to forecasting in emergent areas of research and gauging opinions about new policies. This approach has also been applied extensively as a heuristic in health care management [25,26]. Experts are invited, in 3 rounds of polls, to give their anonymous opinions on a topic. Through an

iterative process, the goal is to establish consensus opinions across the group.

Employing a purposive sampling methodology, the research team compiled a list of 92 participants with expertise in open notes in mental health. There is no universal agreement about the sample size for Delphi polls [27]; however, following previous surveys, our aim is to maximize the volume of responses balanced against maintaining high response rates between surveys [28,29]. As the survey was administered during the COVID-19 pandemic, it was uncertain how many responses we might obtain, and this factored into the decision to invite as many suitable experts as possible. The list was compiled after joint meetings in which the research team examined published research, gray literature, mass media articles, and personal connections to derive as inclusive a list as possible. Acknowledging the challenges associated with defining expertise in a given domain, we interpreted expertise as individuals who had experience, as clinicians, of sharing mental health notes with patients; patients with mental health diagnoses, including patient advocates and peers who had first-hand experience or knowledge of the practice; chief information officers, chief medical information officers, or directors of divisions of health organizations who had implemented sharing mental health notes; and informaticians and other health researchers, including patient researchers, who had published significant contributions within the field of open notes.

To ensure an international perspective, we specifically invited individuals from countries and health systems where clinical note sharing has been implemented. Measures were also taken to ensure gender, age, and demographic diversity. The study received ethical approval from the Beth Israel Deaconess Medical Center Institutional Review Board in April 2020 (reference number 2020P000218) and the University of Plymouth, United Kingdom. Invited participants were advised that the survey was confidential, and their identity would be restricted to a key member of the research team (AK). All the respondents provided informed consent before participating.

Prospective panelists were contacted via email in August 2020 with an invitation and internet link to the survey. Invitees were also informed that participation was voluntary, unpaid, and confidential to the survey team. Participants' names were replaced with a study ID number by AK to preserve anonymity during data analysis.

The Questionnaire

We created an electronic survey using JISC Online Surveys hosted by the University of Plymouth, United Kingdom [30]. The survey was conducted in English. Participants were sent 3 reminders, 1 week apart, and were given 4 weeks to respond to each round. The first round comprised questions about demographic information and the nature of participants' expertise with open notes in mental health. This was followed by four sections with a total of 6 open-ended questions on sharing mental health notes and an additional open-ended question allowing participants to comment on the survey or submit additional responses (Textbox 1; Multimedia Appendix 1). The sections comprised (1) Effects on patients (2 open-ended questions), (2) Effects on clinicians (1 open-ended question), (3) Training and education (2 open-ended questions), and (4) Policy regulations (1 open-ended question). Responses to section 1 were used to form 2 additional rounds of the Delphi survey, and the results will be published elsewhere. In this study, we focused only on participants' open-comment responses to sections 3 and 4, along with the response to the additional open-ended question.

Textbox 1. Round 1 open-ended questions.

Effects on patients

- What, in your opinion, are the benefits, if any, of sharing mental health notes with patients?
- What, in your opinion, are the harms, if any, of sharing mental health notes with patients?

Effects on clinicians

- What, in your opinion, are the effects, if any, on clinicians of sharing mental health notes with their patients?

Training and education

- Should mental health clinicians be trained on how to write clinical notes for patients? If so, what should such training encompass?
- Should mental health patients receive guidance on how to read their mental health notes? If so, what should such guidance encompass?

Policy regulations

- What policy regulations, if any, should be in place for patient access to mental health notes?

Comments

- Do you have any other comments about sharing online mental health notes with patients?

Qualitative Survey Component

Descriptive content analysis was used to investigate the responses using the following steps [31,32]. First, transcripts were read by CB, MH, and JT to familiarize themselves with responses. Second, a process was employed in which brief descriptive labels (*codes*) were applied to comments; for some comments, multiple codes were applied. This widely used

method is considered an efficient methodology for qualitative data analysis [32]. Comments and codes were reviewed by CB, MH, and JT, with revisions leading to further refinement of the codes. Subsequently, first-order codes were grouped into second-order themes based on the commonality of meaning. CB, MH, and JT reviewed and refined the final themes.

Results

Overview

From a total of 92 experts from 6 different countries, 76% (70) of them responded. Among the 70 respondents, 50% (35) were female, and the mean age was 50 years (SD 11.52 years; Table 1). Of the 70 participants, 34% (24) had a PhD degree, and 64% (45) were currently working in clinical practice (Table 2). The mean number of years of experience working in health care, in open notes research, or as a patient advocate was 16 years. All respondents left comments (10,445 words), which were typically brief (1 phrase or 2 sentences).

As a result of the iterative process of content analysis, participants' comments on clinician training, patient guidelines, policy recommendations, and "any other comments" yielded distinctive themes. Owing to the limitations of the data set, these emergent themes reflected the topics of the questions and included (1) clarity about provider policies on exemptions, (2) providing patients with basic information about open notes, (3) clinician training in writing mental health notes, and (4) managing patient-clinician disagreements about mental health notes (Figure 1). These themes were further subdivided into categories, which are described below with illustrative comments.

Table 1. Demographic characteristics of respondents in round 1 (n=70).

Characteristic	Value[a]
Gender, n (%)	
Female	35 (50)
Male (including transgender male)	35 (50)
Age (years)	
Mean (SD)	49.87 (11.52)
Age group (years), n (%)	
20-29	1 (1)
30-39	19 (27)
40-49	14 (20)
50-59	19 (27)
60+	17 (24)
Ethnicity, n (%)	
Asian	6 (9)
Black, African or Caribbean	1 (1)
White	59 (84)
Other	2 (3)
Declined to answer	2 (3)
Country of residence, n (%)	
Canada	3 (4)
Estonia	1 (1)
Norway	3 (4)
Sweden	12 (17)
United Kingdom	4 (6)
United States	47 (67)

[a]Owing to rounding off, not all percentages may add to the total.

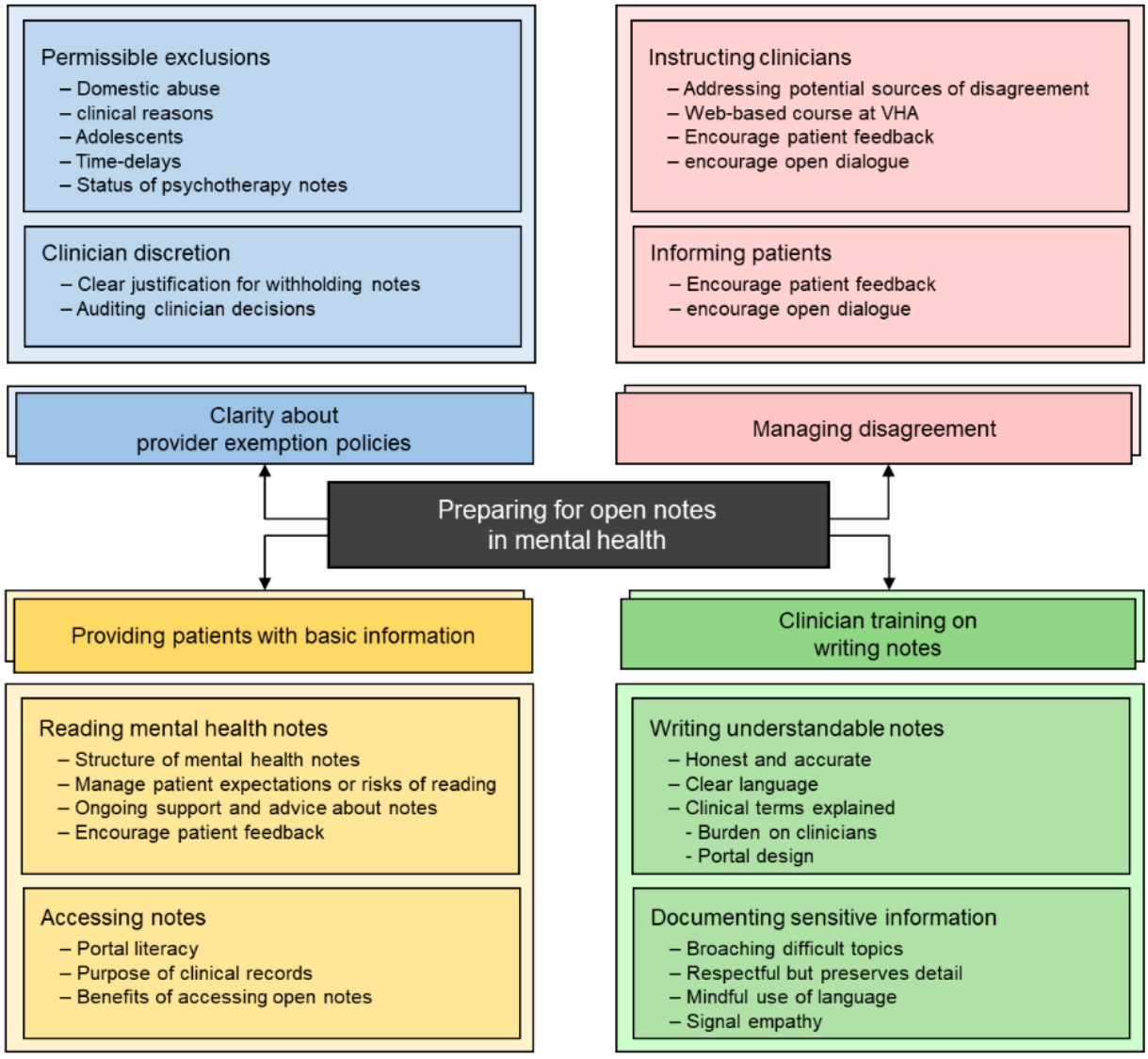
Table 2. Expertise of respondents in round 1 (n=70).

Characteristic	Value[a]
PhD degree, n (%)	24 (34)
Biochemistry, cell biology or molecular pharmacology	3 (4)
Informatics (including eHealth, technology and health engineering)	9 (13)
Psychology (including biological and clinical)	8 (11)
Other (including economics, mathematics, medicine, and philosophy)	4 (6)
Works in a clinical practice, n (%)	45 (64)
Psychiatry (including adult, adolescent and child)	15 (21)
Primary care	10 (14)
Clinical psychology or psychotherapy	7 (10)
Psychiatric nursing	4 (6)
Pediatrics	3 (4)
Social work	3 (4)
Palliative care or home hospice	2 (3)
Peer support	2 (3)
Hospitalist	1 (1)
Radiology	1 (1)
Experience of working in health care, open notes, or patient advocacy (years), mean (SD)	16.32 (12.23)
Occupation or expertise related to health, open notes or patient advocacy, n (%)[b]	
Clinician	46 (66)
Researcher	25 (31)
Chief information officer or portal director or medical director	14 (20)
Patient advocate or person with lived experience	5 (7)
Social worker	1 (1)

[a]Owing to rounding off, not all percentages may add to the total.

[b]Some participants indicated more than one area of expertise.

Figure 1. Themes and categories. VHA: Veterans Health Administration.



Clarity About Provider Policies on Exemptions

Permitted Exclusions

Participants frequently expressed the view that exclusions to note sharing should be permissible in circumstances where access might lead to patient harm. Many comments referred to domestic abuse situations, for example, "notes that could endanger patients if read by eg, a relative/spouse." As one participant noted:

> *If the relative forces the patient to log in to the portal, just the sight of mental health notes/visits might raise suspicion that the patient has been "gossiping."*

The view that exclusions should be permitted if there was "a specific clinical reason" was also commonly voiced, and participants offered suggestions for when hiding notes might be appropriate. One participant suggested that exemptions from sharing should be permitted among "patients with diagnoses or conditions that they believe will be destabilizing for the patient." Other respondents offered more specific clinical scenarios or conditions, including "sexual trauma notes," persons with "delusional" symptomatology, those with "multiple personality disorder, bipolar types I & II, schizophrenia, and active suicidal ideation," and persons with "psychosis, personality disorders, and substance abuse disorders."

In other cases, such as access to notes by adolescents, a few participants expressed reservations or uncertainties about suitable policies and restrictions. As one expert commented:

> *What should be done with adolescents and their parents' access to these notes?*

Offering a different perspective, a few participants proposed the broader view that there should be, "exclusion of all psychotherapy notes" from mandatory sharing. However, several respondents remained neutral about opening mental health notes, asserting, "we need to do some more research," and "the data needs to be collected to document the reality of benefits and impacts."

Expanding on the notion of patient exclusions, some participants proposed that time delays to sharing mental health notes should

be "temporarily possible" or that "therapy notes may need to be held longer before sharing." One participant suggested:

> *There should be a delay period for mental health notes of at least a week.*

Drawing on his or her experience, another respondent noted:

> *We have 12 years of experience of sharing clinical notes with mental health patients in Estonia. We use [an] opt-out concept in [our] nation-wide health information system. The law gives the physician the right to close patient data from the patient for up to 6 months. After that the notes should be disclosed.*

Notably, however, not all participants agreed that exemptions to sharing mental health notes should be permitted and that "open notes means open notes." Some comments reflected the view that restricting access would be unethical, for example:

> *I do not feel that it is ethically correct to set restrictions based solely on diagnoses.*
>
> *Unilateral paternalistic decisions about what should or should not be disclosed to patients is not acceptable.*
>
> *Anything short of real-time patient access and patient control of how and when they wish to access their own record demonstrates an inequity in patient-provider partnerships.*

Clinician Discretion

Addressing the issue about when patient exclusions to opening mental health notes might apply, many comments conveyed the view that "clinicians should be given discretionary power to withhold part or all of the notes." As another respondent noted:

> *I would advocate for clinician judgment always having the final authority to revoke or block notes from patients in very specific instances.*

One respondent dissented from this position, noting:

> *I don't like the idea of limiting the notes to certain patients because we then get into tricky ethical areas and subjectivity on the part of the clinician.*

Elaborating further on the idea of discretionary judgments, however, some participants emphasized that additional measures must be in place. These comments emphasized that, with the responsibility of making decisions about when to block notes, the burden of justification for doing so should be on clinicians. For example:

> *The onus should be on the clinician to record why access should be restricted, giving a review date for when this restricted access should be reconsidered.*

Several participants suggested that formalized checks and balances should be in place to confirm clinicians' decisions about when to deny access to notes. For example:

> *there need[s] to be at least two clinicians signing off on that exclusion before it can be invoked.*

One participant proposed the need for "regular audits and monitoring...to increase transparency of which patient records are blocked and why." Going further, another participant suggested:

> *Organizations should be required to perform audits, by a panel of clinicians and paid service users, at intervals on: (a) the proportion of patients that are accessing their mental health records and organizations with very low rates of access should have their processes for access assessed; (b) the proportion of records that have restricted access and how much of that restricted access is past its "sell by date."*

Providing Patients With Basic Information About Open Notes

Accessing the Notes

Basic issues such as patients' digital literacy and their knowledge of using health portals formed an emergent category. For example:

> *[P]atients should receive information about how to actually reach the notes (the technical part, how to access the information).*

Respondents also emphasized the need to communicate basic information about medical records; as one participant wrote:

> *Explain the purpose of the health record and mental health note.*

Several comments also suggested that patients should be informed about the potential value of open notes, including "how it can lead to better outcomes" and how reading their notes can "inform treatment engagement and decisions."

Reading Mental Health Notes

Multiple comments emphasized the importance of informing patients about how the notes were structured. Some participants suggested that such guidance should include "being given basic information about the mental state examination." Another participant noted:

> *The mental status exam has caused some consternation amongst patients until it is explained to them.*

Another common subcategory was preparing patients to manage their expectations about the content of notes, including "potential areas of friction or frustration;" "why a diagnosis is needed, and that this diagnosis may be fluid;" and reassuring patients that, "the note/diagnoses is not judgment (especially for substance use and personality disorders)." Many participants also described the importance of informing patients about the risk of being upset by what they read, for example:

> *Providing education about the potential emotional response to notes is critical.*

Several respondents suggested that patients should be advised that they will be supported if they choose to read their notes. As one patient advocate noted:

> *I was a patient in a psychiatric ward for 20 years. I recently obtained all my notes. It is fascinating and also disturbing to see what professionals thought and*

> *said about me. It is important for patients to be supported while reading their notes, to prevent relapse or trauma.*

Participants frequently noted that patients would need advice on how to raise questions about what they read in their notes. Some comments emphasized that patients should be "encouraged to bring questions" about their notes to clinicians.

Clinician Training on Writing Mental Health Notes

Writing Understandable Notes

Participants commented on the need to train clinicians on "how to write so that a layman may understand the notes." Multiple comments suggested that clinicians should also be advised on the "degree of medical jargon" that might be documented. Some respondents suggested that there should be a "reduction in jargon" or "avoidance of confusing medical jargon and abbreviations." One participant went further by suggesting that clinicians should also become knowledgeable about common patient terminology:

> *There is an ever-growing list of colloquial expressions and a lexicon for describing diagnoses and symptoms that exists outside of the medical community. And sometimes it's the same word a clinician would use but it means something totally different on a Reddit mental health forum.*

Exploring another aspect of this category, multiple comments expressed the importance of training to preserve details in clinical documentation. Some stated that clinicians should not "dumb down" their notes that needed to be "accurate, objective, truthful." One participant highlighted the need to instruct clinicians on the "documentation of uncertainty."

In general, most comments suggested that the onus should be on clinicians to modify documentation practices, for example:

> *Clinicians should always be writing on the assumption that their audience has zero training...Effective communication is necessary for the job.*

However, one respondent presented a portal design solution:

> *To solve the conflict between professional language precision, and lay person comprehension, one practical solution may be to embed a dictionary in the journal [medical records]. I have seen a proof of concept of such system. The patient was presented a journal text, and by hovering the mouse over a medical record, a dictionary box was presented.*

Documenting Sensitive Information

One frequently identified predicament was the need for guidance on "how to describe sensitive matters" and "difficult topics," including "how to manage situations/write content that may potentially be perceived in a negative light in a way that is clinically appropriate and respectful." Some respondents emphasized that clinicians would need to be instructed that "important information should not be left out due to fear about the patient's reaction." As one participant noted:

> *a patient may discuss a sensitive issue that they do not want in the record and the clinician needs to know how to document information in a way that the patient is comfortable having in writing.*

More generally, participants recommended that clinicians should be trained to adopt a "mindful approach" in documenting mental health notes and to use "patient-friendly" language. Respondents frequently emphasized that clinicians would require training in writing notes that were "not perceived as demeaning," in adopting "less inflammatory terminology," and in choosing language "to avoid stigma or embarrassment." As one respondent noted:

> *[M]any patients stumble over "patient complains of" or "affect" or other common psych mental health terms.*

One respondent suggested that it would be helpful for clinicians to be provided with:

> *[a] list of words that appear to be judgmental or offensive that are frequently used (eg, patient lied, patient was aggressive, patient denied, patient is non-compliant, patient was upset).*

Participants also emphasized the need for training in writing notes with *empathy* and *warmth*. Some comments highlighted an opportunity to use notes to provide greater patient engagement and motivation for treatment goals. As one participant noted, training should also encompass "a framework that acknowledges strengths and doesn't just pathologize;" or as another commented, "education should include an examination of patient strengths as well as deficits; too often MH [mental health] notes can be of the deficit model."

Managing Patient-Clinician Disagreement About Mental Health Notes

Instructing Clinicians

Another category of comments related to clinician training addressed potential disagreements arising from patient access to their mental health notes. As one participant remarked:

> *Clinicians should receive training on the ethics of sharing notes with patients and how to address conflicts.*

Other comments suggested that training should encompass advice about how to discuss disagreements *in vivo* during visits, for example:

> *Such training is not about writing notes, but rather about how to communicate with patients about their illnesses in a way that promotes shared understanding and points of open disagreement (Rather than clandestine documentation).*

Offering ideas on how to prepare clinicians for practice dilemmas, participants suggested the need to "provide tips, templates and scripts on how to address certain situations so clinicians feel more confident." Several respondents cited "the web-based course from VHA [Veteran's Health Administration]" as "a useful starting point" (a webinar on how

to write mental health notes [20]). One participant noted the importance of patient feedback to improve writing notes:

> *when mental health clinicians are first learning how to write notes they should have patient evaluators. For more seasoned clinicians, a modified version of this training could take place.*

Soliciting Patient Collaboration

Multiple comments also described the importance of providing advice to patients about what to do if they detected inaccuracies or omissions in their notes, including "how to handle anything they feel is factually incorrect or was misunderstood in terms of content." Recommendations also encompassed perceived discrepancies in psychotherapy notes, for example, explaining to the patient that, "possible mismatches can always be sorted out during the next session." Several participants noted the positive dimension of soliciting patient feedback. Some comments suggested that patients should be counseled that any errors or feedback present opportunities for working "collaboratively" and increasing "open dialogue between patient and provider."

Discussion

Principal Findings

This qualitative study provides foundational cross-cultural insight into the views of a range of experts on open notes in mental health care. We identified 4 major categories related to opening mental health notes: (1) clarity about provider policies on exemptions, (2) providing patients with basic information about open notes, (3) clinician training in writing notes, and (4) managing patient-clinician disagreements.

Clarity about provider policies included both views that exclusions to note sharing should be permissible under certain circumstances and views that clinicians should be given discretionary power to decide when notes should be shared or not. However, these views were not fully agreed upon, as some participants stressed that restricting access would compromise clinicians' ethical duties. Many respondents discussed special handling of notes around certain mental health conditions such as "psychosis, personality disorders, and substance abuse disorders," among others. Although there is research on open notes and mental health, these conditions are often excluded in studies, resulting in a lack of knowledge. Patients with these disorders may react differently to mental health notes, but separating stigma and assumptions from facts is critical. For example, patients with psychosis are active users of mental health apps and are not paranoid about telehealth, despite common assumptions about their use of technology [33-35]. Open notes could make patients with psychosis paranoid; however, the opposite effect may be achieved. Immediate research is necessary to guide impending implementation efforts. As participants suggested, formalized monitoring is needed to confirm clinicians' decisions and ensure that patients are not wrongfully denied access to their notes.

Although recent research suggests that mental health patients seek the same features on health system portals as medical patients [36], it must be acknowledged that people with mental health conditions, such as schizophrenia, are likely to have less access to portals and, thus, less ability to partake in open notes. At the same time, internet access through smartphones is increasing globally, and a recent study from Sweden indicated that 77% (n=11,001,189) of the visits to the national patient portal were made from a smartphone [37]. Ensuring easy mobile access to patient portals is an important means of increasing accessibility.

Supporting patients on how to access and read their mental health notes was a topic of less controversy. Participants suggested that patients may need basic information on how to access patient portals and notes and on the purpose and content of a health record and mental health notes. The proposed guidance included ways of preparing patients for what to expect when reading their mental health notes to reduce the risk of misunderstandings or harmful emotional responses. Similarly, the need for clinician training on writing notes was also agreed upon by most participants, including concrete advice on how to write more understandable notes and how to address sensitive topics. Participants agreed that information should not be left out of the record for fear of causing patient distress; however, clinicians should be guided on how to address these topics both in the conversation with the patient and in the note.

Disagreements between patients and clinicians caused by the notes were also described as an issue requiring more than just training in how to write better notes. Rather, notes should be seen as one component in the overall communication with the patient, and clinicians should be supported in how to address, and hopefully avoid, disagreement and conflict with patients both in the visit and through written communication. Patient feedback on note writing was highlighted as an important tool. In addition, participants also suggested providing patients with instructions or guidance on what to do when in disagreement with a note to facilitate dialogue and collaboration rather than conflict.

Finally, a topic that was not raised by the respondents in this study was differential diagnoses. In a recent survey of US physicians with experience of open notes, approximately 23% (n=176) of physicians reported changing how they wrote differential diagnoses [38]. The omission of answers focusing on this topic by our respondents could indicate that this is less of a consideration in mental health care. However, in a Swedish survey study, 1 in 5 (22%, n=147) mental health clinicians [8] admitted writing less candid notes, and in a US Veteran's Health Administration study, the majority (69%, n=108) of mental health clinicians reported writing fewer details [9]. That open notes may have an impact on how notes are written seems clear, but further research is needed to further understand the types of changes and their consequences [39].

Strengths and Limitations

This study provides a foundational qualitative exploration of expert opinions on open notes in mental health. Importantly, this inquiry builds on previous survey research conducted in this area by focusing on expert opinions, for the first time, on pressing questions about policy, clinician training, and patient guidance. The survey benefits from a wide range of expertise drawn from countries and health organizations where patients

have access to their mental health notes. The international diversity and breadth of expertise of the respondents from 6 countries are major strengths of the survey. For web-based surveys, a 50% response rate for is considered high [40]; our survey secured a response rate of 76%, which was another major strength of the study.

This study has several limitations. Comments were brief—only 1 or 2 sentences or written in bullet points—restricting a more in-depth understanding of participants' views. In addition, owing to the limitations of web-based surveys, it was not possible to probe or explore respondents' comments to obtain a more in-depth understanding of their views. Although 11% (8/70) of the respondents identified as patients or patient advocates, we suggest future research should directly solicit the views of mental health patients as experts, particularly those with serious mental illness, who work outside of professional health care and academia.

In addition, most of our respondents were White and well educated, which may have affected their responses. Perhaps reflecting on this, questions about access to patient portals received fewer comments. Although this may have reflected the focus of the survey on mental health, many disadvantaged social groups (older patients, those with fewer years of formal education, and vulnerable patient populations) risk losing out on the benefits of access to their clinical notes [41]. For example, in the United States, not everyone has internet access or experience in the use of digital devices to be able to log on to health system portals or read their notes [42].

To address these limitations, further focus groups or interviews would help to facilitate a richer and more nuanced exploration of patients' and mental health clinicians' perspectives on the themes raised in this survey, and further qualitative research is warranted. Future research could aim to better understand the root causes of clinician omission of information, less candid notes, and writing fewer details. For example, there may be personal apprehension and fear by clinicians of increased workload, appointment duration, and/or being questioned or confronted. Such surveys or interviews might address clinician rationales that may not be widely or openly disclosed.

Conclusions

The results of this study indicate that there is a need for training and support for both patients and clinicians regarding the practice of sharing mental health notes and that clear policies are needed to guide clinicians in the process of sharing and to ensure that patients are given access to their information.

We observe a tension in the results between complete transparency and clinicians' need to be able to exclude certain information to prevent patient harm. Although participants generally favored note sharing in mental health care, some stressed the clinicians' autonomy to make fine-grained decisions regarding what information to share, whereas others stressed patients' rights to access all information and the ethical risks of leaving the decision to share to the individual clinician. Evidently, there is a need for evidence-based policies in this area. Surprisingly, few participants raised issues regarding the digital divide and the risk that some patients may not be able to access their notes, possibly indicating that there is still more focus on challenges regarding those patients who *do* access their notes.

The results of this study highlight the need for further training and support for both clinicians and patients regarding note sharing and more thoughtful refinements to user-friendly portal design. A major part of such training and support must address the need for culture change and a shift in mindset toward a more collaborative approach to patient-provider relationships. We suggest that priority should be on training clinicians, including students, on how to write mental health notes. There is a pressing need for transparent systems that support flexible and safe sharing of notes in mental health care. Engaging both patient advocates and eHealth design teams in this study is essential to forge innovative strategies that enable patient understanding of medical terms and feedback on notes, as suggested by the respondents. We hope that this study will provide both direction and inspiration for further research and policy reflection on patient access to their mental health notes.

Ethical Approval

This study was deemed exempt by the Beth Israel Deaconess Medical Center Institutional Review Board on April 10, 2020 (reference number 2020P000218), and the University of Plymouth, United Kingdom.

Acknowledgments

The authors would like to thank the experts who participated in this survey. The authors would also like to thank Dr John Santa for his assistance with administering the survey and for offering valuable feedback on a previous draft of the manuscript. CB was supported by a John B. Keane Fellowship. The study was also supported by FORTE—the Swedish Research Council for Health, Working Life and Welfare through the research projects "PACESS" (2016-00623) and "Beyond Implementation of eHealth" (2020-01229).

Authors' Contributions

CB conceived the project. CB, AK, JT, SO, MH, and CMD designed the survey. AK, CB, KH, SO, DW, JT, LS, and MH administered the survey. CB, AK, JT, and MH analyzed the results. CB wrote the first draft. AK and CB developed figures and tables. AK, CB, CMD, SO, DW, LS, and MH contributed to the revisions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Round 1 survey.
[PDF File (Adobe PDF File), 1897 KB - mental_v8i4e27397_app1.pdf]

References

1. Essén A, Scandurra I, Gerrits R, Humphrey G, Johansen MA, Kierkegaard P, et al. Patient access to electronic health records: differences across ten countries. Health Policy and Technol 2018;7(1):44-56. [doi: 10.1016/j.hlpt.2017.11.003]
2. Björgvinsson E, Ehn P, Hillgren PA. Participatory design and "democratizing innovation". In: Proceedings of the 11th Biennial Participatory Design Conference. 2010 Presented at: PDC '10: The 11th Biennial Participatory Design Conference; November, 2010; Sydney, Australia. [doi: 10.1145/1900441.1900448]
3. Gregory J. Scandinavian approaches to participatory design. Int J Eng Educ. 2003. URL: https://tinyurl.com/57p8zskv [accessed 2021-03-24]
4. Health and Human Services Department. 21st Century Cures Act: interoperability, information blocking, and the ONC health IT certification program. Federal Register. 2020 May 1. URL: https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification [accessed 2020-07-15]
5. Blease C, Walker J, DesRoches CM, Delbanco T. New U.S. law mandates access to clinical notes: implications for patients and clinicians. Ann Intern Med 2021 Jan;174(1):101-102. [doi: 10.7326/m20-5370]
6. Rucker DW. Implementing the Cures Act — bringing consumer computing to health care. N Engl J Med 2020 May 07;382(19):1779-1781. [doi: 10.1056/nejmp2003890]
7. Kristiansen E, Johansen M, Zanaboni P. Healthcare personnels' experience with patients' online access to electronic health records: differences between professions, regions, and somatic and psychiatric healthcare. In: Proceedings of the 17th Scandinavian Conference on Health Informatics. 2019 Presented at: SHI 2019: 17th Scandinavian Conference on Health Informatics; November 12-13, 2019; Oslo, Norway p. 93-98 URL: https://ep.liu.se/en/conference-article.aspx?series=ecp&issue=161&Article_No=16
8. Petersson L, Erlingsdóttir G. Open notes in Swedish psychiatric care (Part 2): survey among psychiatric care professionals. JMIR Ment Health 2018 Jun 21;5(2) [FREE Full text] [doi: 10.2196/10521] [Medline: 29929946]
9. Dobscha SK, Denneson LM, Jacobson LE, Williams HB, Cromer R, Woods S. VA mental health clinician experiences and attitudes toward OpenNotes. Gen Hosp Psychiatry 2016 Jan;38:89-93. [doi: 10.1016/j.genhosppsych.2015.08.001] [Medline: 26380876]
10. Klein JW, Peacock S, Tsui JI, O'Neill SF, DesRoches CM, Elmore JG. Perceptions of primary care notes by patients with mental health diagnoses. Ann Fam Med 2018 Jul 09;16(4):343-345 [FREE Full text] [doi: 10.1370/afm.2287] [Medline: 29987083]
11. Peck P, Torous J, Shanahan M, Fossa A, Greenberg W. Patient access to electronic psychiatric records: a pilot study. Health Policy Technol 2017;6(3):309-315. [doi: 10.1016/j.hlpt.2017.06.003]
12. Denneson LM, Chen JI, Pisciotta M, Tuepker A, Dobscha SK. Patients' positive and negative responses to reading mental health clinical notes online. Psychiatr Serv 2018 May 01;69(5):593-596. [doi: 10.1176/appi.ps.201700353] [Medline: 29493408]
13. Cromer R, Denneson LM, Pisciotta M, Williams H, Woods S, Dobscha SK. Trust in mental health clinicians among patients who access clinical notes online. Psychiatr Serv 2017 May 01;68(5):520-523 [FREE Full text] [doi: 10.1176/appi.ps.201600168] [Medline: 28142383]
14. Denneson LM, Pisciotta M, Hooker ER, Trevino A, Dobscha SK. Impacts of a web-based educational program for veterans who read their mental health notes online. J Am Med Inform Assoc 2019 Jan 01;26(1):3-8 [FREE Full text] [doi: 10.1093/jamia/ocy134] [Medline: 30445648]
15. Moll J, Rexhepi H, Cajander Å, Grünloh C, Huvila I, Hägglund M, et al. Patients' experiences of accessing their electronic health records: national patient survey in Sweden. J Med Internet Res 2018 Nov 01;20(11):e278 [FREE Full text] [doi: 10.2196/jmir.9492] [Medline: 30389647]
16. Walker J, Leveille S, Bell S, Chimowitz H, Dong Z, Elmore JG, et al. Correction: Opennotes after 7 years: patient experiences with ongoing access to their clinicians' outpatient visit notes. J Med Internet Res 2020 Apr 30;22(4) [FREE Full text] [doi: 10.2196/18639] [Medline: 32352924]
17. Petersson L, Erlingsdóttir G. Open notes in Swedish psychiatric care (Part 1): survey among psychiatric care professionals. JMIR Ment Health 2018 Feb 02;5(1):e11 [FREE Full text] [doi: 10.2196/mental.9140] [Medline: 29396386]
18. Strudwick G, Yeung A, Gratzer D. Easy access, difficult consequences? Providing psychiatric patients with access to their health records electronically. Front Psychiatry 2019 Dec 16;10:917 [FREE Full text] [doi: 10.3389/fpsyt.2019.00917] [Medline: 31920763]

19. O'Neill S, Chimowitz H, Leveille S, Walker J. Embracing the new age of transparency: mental health patients reading their psychotherapy notes online. J Ment Health 2019 Oct 31;28(5):527-535. [doi: 10.1080/09638237.2019.1644490] [Medline: 31364902]
20. Dobscha SK, Kenyon EA, Pisciotta MK, Niederhausen M, Woods S, Denneson LM. Impacts of a web-based course on mental health clinicians' attitudes and communication behaviors related to use of OpenNotes. Psychiatr Serv 2019 Jun 01;70(6):474-479. [doi: 10.1176/appi.ps.201800416] [Medline: 30890047]
21. Blease C, Torous J. Opening mental health notes: 7 tips to prepare clinicians. Psychol Today. 2020. URL: https://www.psychologytoday.com/us/blog/digital-mental-health/202010/opening-mental-health-notes-7-tips-prepare-clinicians [accessed 2020-12-11]
22. Blease CR, O'Neill S, Walker J, Hägglund M, Torous J. Sharing notes with mental health patients: balancing risks with respect. Lancet Psychiatry 2020 Nov;7(11):924-925 [FREE Full text] [doi: 10.1016/s2215-0366(20)30032-8]
23. Blease C. Are mental health patients entitled to see their medical notes? Psyche. 2020. URL: https://psyche.co/ideas/are-mental-health-patients-entitled-to-see-their-medical-notes [accessed 2020-08-26]
24. Blease CR, Walker J, Torous J, O'Neill S. Sharing clinical notes in psychotherapy: a new tool to strengthen patient autonomy. Front Psychiatry 2020 Oct 28;11. [doi: 10.3389/fpsyt.2020.527872]
25. Donohoe H, Stellefson M, Tennant B. Advantages and limitations of the e-Delphi technique. Am J Health Educ 2013 Jan 23;43(1):38-46. [doi: 10.1080/19325037.2012.10599216]
26. Custer RL, Scarcella JA, Stewart BR. The modified Delphi technique - a rotational modification. J Voc Tech Educ 1999:50-58. [doi: 10.21061/jcte.v15i2.702]
27. Rubinson L, Neutens JJ. Research Techniques for the Health Sciences. New York: Macmillan Pub Co; 1987.
28. Norcross JC, Pfund RA, Prochaska JO. Psychotherapy in 2022: a Delphi poll on its future. Prof Psychol Res Pr 2013 Oct;44(5):363-370. [doi: 10.1037/a0034633]
29. Blease C, Kharko A, Locher C, DesRoches CM, Mandl KD. US primary care in 2029: a Delphi survey on the impact of machine learning. PLoS One 2020;15(10) [FREE Full text] [doi: 10.1371/journal.pone.0239947] [Medline: 33031411]
30. Jisc online surveys. URL: http://www.onlinesurveys.ac.uk [accessed 2021-04-08]
31. Mayring P. Qualitative content analysis: theoretical foundation, basic procedures and software solution. SSOAR Open Access Respository. 2014. URL: https://www.ssoar.info/ssoar/bitstream/handle/document/39517/ssoar-2014-mayring-Qualitative_content_analysis_theoretical_foundation.pdf [accessed 2021-03-24]
32. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
33. Firth J, Torous J. Smartphone apps for schizophrenia: a systematic review. JMIR Mhealth Uhealth 2015 Nov 06;3(4):e102 [FREE Full text] [doi: 10.2196/mhealth.4930] [Medline: 26546039]
34. Torous J, Wisniewski H, Liu G, Keshavan M. Mental health mobile phone app usage, concerns, and benefits among psychiatric outpatients: comparative survey study. JMIR Ment Health 2018 Nov 16;5(4) [FREE Full text] [doi: 10.2196/11715] [Medline: 30446484]
35. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar 22;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
36. Strudwick G, Booth RG, McLean D, Leung K, Rossetti S, McCann M, et al. Identifying indicators of meaningful patient portal use by psychiatric populations. Inform Health Soc Care 2020 Oct 01;45(4):396-409. [doi: 10.1080/17538157.2020.1776291] [Medline: 32603617]
37. Hägglund M, Blease C, Scandurra I. Mobile access and adoption of the Swedish national patient portal. Stud Health Technol Inform 2020 Nov 23;275:82-86. [doi: 10.3233/SHTI200699] [Medline: 33227745]
38. DesRoches CM, Leveille S, Bell SK, Dong ZJ, Elmore JG, Fernandez L, et al. The views and experiences of clinicians sharing medical record notes with patients. JAMA Netw Open 2020 Mar 02;3(3) [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.1753] [Medline: 32219406]
39. Blease C, Torous J, Hägglund M. Does patient access to clinical notes change documentation? Front Public Health 2020 Nov 27;8 [FREE Full text] [doi: 10.3389/fpubh.2020.577896] [Medline: 33330320]
40. Baruch Y, Holtom BC. Survey response rate levels and trends in organizational research. Hum Relat 2008 Aug 01;61(8):1139-1160. [doi: 10.1177/0018726708094863]
41. Blease C, Fernandez L, Bell SK, Delbanco T, DesRoches C. Empowering patients and reducing inequities: is there potential in sharing clinical notes? BMJ Qual Saf 2020 Oct 18;29(10):1-2. [doi: 10.1136/bmjqs-2019-010490] [Medline: 32188711]
42. Rodriguez JA, Lipsitz SR, Lyles CR, Samal L. Association between patient portal use and broadband access: a national evaluation. J Gen Intern Med 2020 Dec 10;35(12):3719-3720. [doi: 10.1007/s11606-020-05633-4] [Medline: 31925739]

Edited by G Strudwick; submitted 23.01.21; peer-reviewed by S Brudnicki, A Hakkim; comments to author 21.02.21; revised version received 28.02.21; accepted 28.02.21; published 16.04.21.

Please cite as:
Blease C, Torous J, Kharko A, DesRoches CM, Harcourt K, O'Neill S, Salmi L, Wachenheim D, Hägglund M
Preparing Patients and Clinicians for Open Notes in Mental Health: Qualitative Inquiry of International Experts
JMIR Ment Health 2021;8(4):e27397
URL: https://mental.jmir.org/2021/4/e27397
doi:10.2196/27397
PMID:33861202

Original Paper

Postsecondary Student Engagement With a Mental Health App and Online Platform (Thought Spot): Qualitative Study of User Experience

Howard W Wong[1,2], MHI; Brian Lo[1,2,3], MHI; Jenny Shi[1], MPH; Elisa Hollenberg[1], MSW; Alexxa Abi-Jaoude[1], MPH; Andrew Johnson[1], BA; Gloria Chaim[4,5,6], MSW; Kristin Cleverley[4,5,7], RN, PhD; Joanna Henderson[4,5,6], PhD; Andrea Levinson[4,8], MD; Janine Robb[8], RN, MSc; Aristotle Voineskos[4,9], MD, PhD; David Wiljer[1,2,3,4], PhD

[1]Office of Education, Centre for Addiction and Mental Health, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]UHN Digital, University Health Network, Toronto, ON, Canada

[4]Department of Psychiatry, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[5]Margaret and Wallace McCain Centre for Child, Youth & Family Mental Health, Centre for Addiction and Mental Health, Toronto, ON, Canada

[6]Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, Toronto, ON, Canada

[7]Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

[8]Health and Wellness, University of Toronto, Toronto, ON, Canada

[9]Slaight Family Centre for Youth in Transition, Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
David Wiljer, PhD
UHN Digital
University Health Network
R. Fraser Elliott Building RFE 3-411
190 Elizabeth Street
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 6322
Email: david.wiljer@uhn.ca

Abstract

Background: There is growing interest in using mobile apps and online tools to support postsecondary student mental health, but most of these solutions have suboptimal user engagement in real-world settings. Poor engagement can limit long-term effectiveness and usefulness of these tools. Previous literature has proposed several theories that link factors such as low usability and poor user-centered design to app disengagement. However, few studies provide direct evidence showing what factors contribute to suboptimal user engagement in the context of mobile mental health apps for postsecondary students.

Objective: This study focuses on understanding postsecondary students' attitudes and behaviors when using Thought Spot, a co-designed mental health app and online platform, to understand factors related to engagement and user experience.

Methods: Students who were given access to Thought Spot for 6 months during a randomized trial of the intervention were invited to participate in one-on-one semistructured interviews. The interviews explored participants' overall experiences and perceptions of the app, along with factors that affected their usage of various features. All interviews were recorded, and template analysis was used to analyze transcripts.

Results: User satisfaction was mixed among users of Thought Spot. The degree of engagement with the app appeared to be affected by factors that can be grouped into 5 themes: (1) Students valued detailed, inclusive, and relevant content; (2) Technical glitches and a lack of integration with other apps affected the overall user experience and satisfaction with the app; (3) Using the app to support peers or family can increase engagement; (4) Crowdsourced information from peers about mental health resources drove user engagement, but was difficult to obtain; and (5) Users often turned to the app when they had an immediate need for mental health information, rather than using it to track mental health information over time.

Conclusions: Content, user experience, user-centeredness, and peer support are important determinants of user engagement with mobile mental health apps among postsecondary students. In this study, participants disengaged when the app did not meet

their expectations on these determinants. Future studies on user engagement should further explore the effectiveness of different features and the relative importance of various criteria for high-quality apps. Further focus on these issues may inform the creation of interventions that increase student engagement and align with their mental health needs.

Trial Registration: ClinicalTrials.gov NCT03412461; https://clinicaltrials.gov/ct2/show/NCT03412461

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.6446

(JMIR Ment Health 2021;8(4):e23447) doi:10.2196/23447

KEYWORDS

transition-aged youth; qualitative study; user experience; help-seeking; mental health; postsecondary; mobile apps; adolescent

Introduction

Facilitating access to mental health support for postsecondary students (youth aged 17-29) is critical to preventing underlying conditions from worsening as these transition-aged youth enter adulthood [1-3]. With widespread ownership of cell phones and a willingness to try web-based services, online and mobile mental health apps have been touted as a promising way to provide mental health information and services to this population [4-7]. Delivering mental health support through mobile apps offers several advantages over traditional in-person services, such as improved ease of access, convenience, lower costs, reduced stigma, and user customization [8,9]. Many digital mental health interventions have been developed and offer a wide range of functions, including self-help, symptom monitoring, cognitive behavioral therapy, and psychoeducation for postsecondary students [8,10,11].

Engagement is often considered when understanding the efficacy of digital behavior change interventions, a category that many digital mental health apps fall under. In the literature, engagement refers to both the subjective experience and behavior when interacting with a digital behavior change intervention [12]. According to traditional computer science and human–computer interface research, engagement as a subjective experience can be characterized as feeling focused, attentive, and satisfied when using a digital technology [12]. In comparison, engagement as a behavior commonly refers to the patterns of usage (eg, frequency, duration, retention rate) and depth of usage (eg, use of a specific app feature) [12-14].

Unfortunately, many mobile mental health interventions that target this population face low user engagement [8,13,15,16]. For example, one study found that among 93 mental health apps on the Google Play store, the median 15-day retention rate was 3.9% [13]. Overall, low engagement can hinder an app's ability to deliver positive outcomes to its users and makes it difficult for researchers to understand the app's long-term efficacy [8,12,13,15,17,18].

Currently, there is a lack of direct evidence to explain low engagement in the postsecondary student population, but a few theories do exist [17,19]. Some theories posit that personal attitudes toward technology and perceptions about seeking help can lead to disengagement. The design, content, and usability of an app have also been proposed as important factors that affect engagement [17,19]. Exploring these theories will provide stakeholders, such as app developers, clinicians, postsecondary institutions, and policymakers, with evidence about the factors that influence student engagement with mobile mental health apps. These investigations can reveal opportunities and drive strategies for increasing student engagement with mobile mental health apps in the future.

This qualitative study seeks to understand user engagement by exploring postsecondary students' experiences on Thought Spot, a mobile mental health mobile app. Thought Spot was created through participatory design research methods and usability testing with college and university students [20,21]. Drawing on the social cognitive theory [22] and the theory of help-seeking [23], the intended purpose of this app is to be an online and mobile resource with features that help students find mental health support, build self-efficacy for seeking help, and increase help-seeking behavior [20,24-26]. For example, Thought Spot contains curated information about mental health and wellness resources for youth, and allows them to geo-locate mental health services. The app also displays mental health services and resources that do not require in-person visits, such as websites, apps, and phone or chat support lines tailored to youth. An in-app crowdsourcing function enables users to add mental health and wellness resources, rate them on a 5-star scale, and write reviews for others to see. Newly added resources appear on a timeline that is updated in real time. Resources are categorized using tagged keywords to allow for easy and intuitive searches. A mood-tracking journal enables users to privately record their thoughts and moods during their help-seeking process and to geo-locate where these thoughts occurred.

The primary objective of this qualitative study is to identify factors that affect postsecondary students' attitudes and behaviors when using the Thought Spot app, and to describe how those factors affected user experience and subjective user engagement.

Methods

Study Design

This qualitative study is part of a larger study that includes a randomized controlled trial (RCT) to evaluate Thought Spot [20,27]. The qualitative approach serves to explore questions that complement the larger study, including those that examine factors affecting adoption of and user engagement with Thought Spot for seeking mental health support. It describes the who, what, and where of participant experiences [28]. This study is also a continuation of the participatory design research process and a progression from prior work on Thought Spot [20,24,25]. In contrast to previous qualitative evaluations of Thought Spot

in theoretical situations, which were emulated during co-design workshops and usability tests, this study focuses on the experiences of youth participants who used the app in their day-to-day lives during the RCT [20,24,25,27].

Recruitment, Sampling, and Participants

Participants were recruited to the qualitative phase of the study through purposeful sampling of a subset of students from the intervention group who completed the Thought Spot RCT [27]. Students who indicated in an end-of-study usability survey that they were interested in participating were identified as potential participants for the interviews. The survey explained that the study would involve a 30- to 60-minute in-person or telephone interview about the user's experience with Thought Spot, and that compensation would be provided. A purposive sampling strategy determined who would be selected for an interview. During the RCT, intervention group participants were sent an email on how to download and access Thought Spot on their personal device [20]. Participants were asked to use the app as needed and they were free to use it in whatever manner they liked. As such, participants were only selected to participate in qualitative interviews if they had logged into the Thought Spot app more than once during the trial between March 2018 and June 2019. User activity was verified by one member of the research team (JS) through a filtered search of Thought Spot's user activity data logs. Second, participants were purposefully sampled in 2 groups because the RCT was a longitudinal study and participants who received Thought Spot started the study at different times. In January 2019, changes were made to Thought Spot to address technical issues, resulting in small differences between the versions used by study participants. Consequently, to obtain a comprehensive picture of user experience and user engagement, participants were grouped into those who finished the Thought Spot trial before January 2019 (Group A) and those who were participating in the trial after that date (Group B). The research team also purposefully sampled a 50:50 split of users with high usage/satisfaction and low usage/dissatisfaction in both groups. Usage and satisfaction were determined from separate analyses of individual-level usage data from the app and participant scores on the adapted Usefulness, Satisfaction, and Ease of Use (USE) questionnaire, respectively [29] (Shi et al, unpublished data, 2021). The adapted questionnaire was part of the end-of-study usability survey administered to all participants in the intervention arm to evaluate usefulness, ease of use, ease of learning, and satisfaction with the Thought Spot app [20,29] (Shi et al, unpublished data, 2021). The full reporting and analysis of usage data and USE questionnaire responses are reported elsewhere (Shi et al, unpublished data, 2021). During the purposeful sampling process of this study, if there were no more participants with high usage/satisfaction or low usage/dissatisfaction left to sample, the research team sampled participants with mixed USE questionnaire scores/usage, such as those with high USE score and low usage or low USE score and high usage. Applying findings from the separate analyses of USE questionnaire responses and usage data, the research team used the median USE Questionnaire Score, 53.78 (IQR 38.89-67.78) out of total score 100, to discern high/low satisfaction users (Shi et al, unpublished data, 2021). The median number of clicks, 14 clicks (IQR 6-22), was used to discern high/low usage users (Shi et al, unpublished data, 2021). Overall, the purposeful sampling criteria were intended to identify a sample of participants with varying degrees of usage and different perceptions about Thought Spot's usability.

Students were invited to participate in an interview through an email that contained a summary of the qualitative study and an informed consent form [14]. Participants were offered an honorarium of CAD $40 (USD $32). Interested individuals submitted a signed consent form and arranged a phone or in-person interview with the research analyst and research trainee (BL and HW, respectively).

The study was approved by the Research Ethics Boards at the Centre for Addiction and Mental Health (REB #023/2017), George Brown College (REB #6004416), Ryerson University (REB # 2017-196-1), and the University of Toronto (REB# 00034725). The study was conducted between February 2019 and August 2019.

Data Collection

After obtaining informed written consent from participants, 2 members of the research team (BL and HW) conducted semistructured interviews using the question guide presented in Multimedia Appendix 1. The interviewers have formal education in health informatics and received training and support from other members of the research team who have expertise in qualitative research. The domains covered in the question guide included general impressions of Thought Spot and its features, the utility and impact of the app for help-seeking and finding resources, how and why the app was used, areas of strength and weaknesses, how using the app related to other help-seeking experiences, and suggestions for how to improve the app. The interviewers used a "funneling" approach for the interview guide [30]. First, they invited participants to share their experience and perspectives on Thought Spot and how they used it to meet their needs. Based on the response, the interviewers probed specific topics, such as usage patterns, changes to the help-seeking process, and what the participant liked or disliked about the app. The interview guide was adjusted iteratively when new patterns emerged during the interviews. Interviews were conducted until the researchers felt that further data collection did not add more depth to the emergent codes or themes [31].

All interviews were audio-recorded, deidentified, and transcribed verbatim by a professional third-party service. Members of the research team checked the transcripts for accuracy and corrected discrepancies.

Data Analysis

The research team used a comparative and iterative thematic approach to developing themes from the interview transcripts. Two authors (HW and BL) analyzed the transcripts to explore themes related to on-app user behavior, motivation for usage, perceptions of the app, and suggested improvements. The analysis followed the procedural steps recommended by Brooks et al [32] for template analysis because this method permits the inclusion of predefined codes. The research team applied steps from directed content analysis to include preliminary codes

from human factors research into the analysis, such as appearance, layout, navigation, and ease of use [33,34].

HW and BL reviewed an initial subset of 3 randomly selected interviews to familiarize themselves with the data before coding. Preliminary coding was then completed on the subset, where keywords were highlighted to guide the development of an initial coding template. Highlighted text from the 3 transcripts was clustered into meaningful codes. Related codes were clustered into a hierarchical structure, with narrow codes organized as subcodes for broader themes. HW and BL created a definition for each theme in the initial coding template and presented them to the rest of the research team, along with exemplary quotes. The research team reviewed the initial coding template and modified it through consensus to ensure representativeness and relevance. The remaining transcripts were coded using the revised coding template. The coding template was revised iteratively, which involved integrating new themes and re-defining existing ones as more transcripts were coded. During coding, members of the research team ensured that all codes relevant to the research question were accounted for in the template. The template was presented to the entire research team at a second session to review and finalize ideas, new themes, and interpretations. Having the broader research team review the template ensured that all perspectives were incorporated in the analysis.

Results

Demographics of Participants

Baseline characteristics of the participants are presented in Table 1. The research team was satisfied that they were not seeing any new data after 17 interviews. A total of 11 interviews were conducted via telephone calls and 6 were conducted in-person; 9 students from Group 1 were interviewed in March 2019 and 8 students from Group 2 were interviewed in June 2019. A total of 13 female (76%) and 4 male (24%) students were interviewed, which reflected the demographics of our RCT study of 481 participants (190/241, 78.8%, identified as female). All 3 academic institutions from the RCT were represented.

Following our template analysis, 5 main themes emerged based on users' experiences and perceptions of the app: (1) Students valued detailed, inclusive, and relevant content on the app; (2) Technical glitches and a lack of integration with other apps affected overall user experience and satisfaction with the app; (3) An app's features can extend beyond the users to support their peers; (4) Crowdsourced information from peers about mental health resources was valuable and sought after, but difficult to obtain; and (5) Users often used the app when they had an immediate need for mental health information, rather than using it to track mental health information over time.

Table 1. Baseline characteristics of participants who were interviewed (N=17).

Characteristic	Value
Gender, n (%)	
Female	13 (76)
Male	4 (24)
Other	0 (0)
Age (years), median (IQR)	23.1 (20.9-25.6)
Interview time (minutes), median (IQR)	32.4 (30.9-34.6)
Classification of participants (USE[a] questionnaire rank and usage data rank), n (%)	
Low satisfaction[b] and low usage[c]	3 (18)
Low satisfaction and high usage	7 (41)
High satisfaction and low usage	3 (18)
High satisfaction and high usage	4 (24)

[a]USE: Usefulness, Satisfaction, and Ease of Use.

[b]Low satisfaction is defined as users with a ≤50th percentile USE score.

[c]Low usage is defined as users with ≤50th percentile number of clicks.

Students Value Detailed, Inclusive, and Relevant Content

Many students believed that getting information about mental health resources was valuable in both the short and long term. Many appeared to use Thought Spot as an information-gathering tool to learn about nearby mental health services. Opinions about the app content appeared to affect user satisfaction. For example, some users were pleased with the level of detail and diversity of mental health information that the app provided (all quotes are presented verbatim, but to improve readability, some filler words such as *like* and *um* have been removed):

> *I liked, when you clicked it, it would tell you what kinds of services it offered. ...There were actually details. ...It showed me what the services were. It showed me what the hours were. It showed me the address. And that's all I really needed at the time. Maybe if I had investigated further, I would have realized that I needed other things, but no. It was good. I was satisfied with it.* [P14]

By contrast, other users were dissatisfied with the content. For example, a few commented on the lack of breadth, depth, and relevant information. These users wanted more services to be available on the app and sought additional details about them.

> *I just have utilized a lot of services...so when I was going in just kind of looking, a bunch of things just weren't included or listed, so I was like, I'm not really sure what people are going to be finding when they come in here...I think that was the only drawback to it, or I guess the piece that it didn't meet was having all of the relevant services listed.* [P08]

Some users were curious about the kinds of support provided by a mental health service, such as whether it offered peer support, professional counseling, or another form of support. There were also suggestions to include more background information about the people providing these services, because it helped users gauge whether a service was inclusive of their needs and preferences.

> *Everybody knows you can go to counselling in your school...but someone who is of colour like might not want to go into that space...so there's a space that's hosted by the student's union that gives peer support but it's not called peer support...having that added into the app could be really good. I found it really helpful and it's very inclusive, that might be a better way for students.* [P07]

There were also other comments related to inclusion, as several users said that the content on the app was too general and did not account for their unique circumstances. For example, some users said that the information on the app did not fully consider factors such as where they lived, where they went to school, whether they had health insurance, and how long it would take to access a service. Consequently, it made it challenging for some participants to understand whether a services or resource was relevant for them.

> *...it didn't really take into account the resources I had as a student...it kind of treated all situations as equals, so let's say, I do have insurance coverage...I know it's better for me to go in to somebody I can pay for and that can see me sooner, but it's not necessarily that the app took that into consideration.* [P02]

> *I go to [School A] and I found a lot of the places were close to [School B] or close to [School C] I just remember that I couldn't use the full app, because I don't live in Toronto and I don't have a lot of time when I'm down there.* [P16]

> *...if I'm like [School C] student...am I really not allowed to use the [other] school's app...it doesn't really divide or split between [the different types of] professionals and students...So some people may think, "Oh, this doesn't apply to me"....* [P13]

Other participants wanted to access more content, tools, and strategies to help them manage their mental health concerns, rather than simply being directed to mental health services or resources. Some wanted more support to be delivered directly through the app.

> *...I have anxiety, so I try to use apps like Calm or different ones that help me calm my anxiety, like Breath, all those different apps...I thought maybe this app was going to help me get those kind of features that the other apps have, like strategies to deal with my mental health issues like anxiety and anxiousness and stress and daily stress, but really it was just providing me with different places to go, I believe, if I understood the purpose of the app properly. So the app wasn't giving me tools, it was just redirecting me.* [P05]

Participants indicated that it was important for content to be comprehensive and relevant. They were satisfied when mental health information details were relevant to their circumstances, needs, and preferences. The lack of relevant details also made it difficult for some users to assess whether content on the app was applicable to them and may have decreased their willingness to use Thought Spot. Furthermore, participants indicated that the app provided sufficient information to get a preliminary and surface-level understanding of what mental health services or resources were offered, but that it lacked the level of detail that some participants needed to motivate them to try the resources or services.

Technical Glitches and a Lack of Integration With Other Apps Affected Overall User Experience and Satisfaction With the App

Several participants identified technical issues as a source of frustration and inconvenience when navigating the app. They described occasional glitches and system lag when using the search and filter feature. This feature was designed to help students make custom searches by selecting key terms to narrow down the services and wellness locations most applicable to them. However, technical issues prolonged the amount of time it took some users to find and retrieve information:

> *...I tried the search. I tried to look at different features that it had...in the first few times I tried it, it was kind of glitchy and I had to go back and restart.* [P05]

> *Sometimes when I went to do something it takes a couple of tries to get the map moving, or if I want to search something, it does take a couple of tries to get it to work, but it doesn't happen all of the time.* [P03]

Integration between Thought Spot and other apps was discussed by several participants. They reported using the app alongside other tools, such as Google and journaling and wellness apps, when looking for mental health resources. However, some wanted an all-inclusive app that could connect them to a variety of these tools directly through the app. The current version of Thought Spot does not integrate with other apps, and some students expressed dissatisfaction with the cumbersome process of switching between several platforms. Some felt that improved integration could create a more seamless experience and increase the likelihood that they would take action after accessing information on Thought Spot:

> *I like to have everything sort of integrated into one application. So given the option, like if I was tracking my fitness app—I once did have a calorie counter and*

> *a meal tracker and my monthly menstrual cycles all in the same app, just because it's too much work to have to go and change apps, and you know, it's more work and I'd be less likely to do it.* [P09]

Despite technical issues and lack of integration, some participants were still intrigued by Thought Spot's potential usefulness for students:

> *...I remember having some trouble with how it operated it on my phone, and that prevented me from using it a lot, I think, during the study, but also, I thought it was a really great idea.* [P15]

Specifically, a few participants liked the app's goal of helping students access mental health support and felt that updating and optimizing the app further would resolve issues related to user experience:

> *...so maybe in the future we have more financial resource to support this application, I think it should be better, and in the long term I think it's really good for students.* [P12]

Participants saw value in the app itself, but the quality of Thought Spot's user experience varied, with some users encountering more issues than others. The main concerns about slow loading and integration with other apps may have interfered with their ability to engage meaningfully and fully with the app.

The App's Functionalities Can Extend Beyond the User to Support Peers

Several participants reported using Thought Spot to help their friends and family members. They described sharing resources and services with others who needed mental health support, even when they did not immediately require services or resources themselves:

> *...I shared the app with my friends, and with some of my friends that have...something they don't want to talk with about to the family or relatives, so I introduced them [to] this app.* [P12]

> *I didn't necessarily go to all of them, but I sent friends to some of the places, like when they needed to go somewhere, I would say, you know, there's this place, it's 100 meters away.* [P09]

The potential of Thought Spot to play a role in providing peer support for mental health concerns was discussed by some participants. They described mental health as a sensitive topic and thought that it was valuable to share information on the app because it might be more trustworthy. One participant suggested adding value by being able to directly share the information with a friend through channels such as social media:

> *Yeah, refer your friends or share it with your friend, because you know what? With this kind of very sensitive issue, sensitive information, you just believe what you trust or believe. So that's why refer your friend, introducing your friend, that should be a function in the app—to share with your friend.* [P12]

Even though some participants did not have a pressing need for mental health help themselves, they used the app to become messengers of mental health–related information within their social circles. In situations where a peer required support, participants explained that Thought Spot gave them information that they could share with their peer and that would help them access the resources or services they needed.

Crowd-Sourced Information From Peers About Mental Health Resources Was a Driver of Engagement, but Was Difficult to Obtain

Thought Spot enables students to crowdsource information, that is, to add mental health services and self-care locations (classified as "spots") and to provide reviews about these resources. Most users agreed that peer reviews were important and valuable because reviews about mental health resources are often difficult to find and reading about other people's experiences can increase motivation to access services:

> *Having the reviews and the comments from peers who have utilized those different groups was...a huge thing that doesn't exist anywhere.* [P08]

The same participant added that evaluating the quality of services was challenging or tricky, and seeing diverse peer reviews gave them a more balanced perspective about a "spot" or resource:

> *And I think through the reviews I'm able to get a little bit more of, kind of a sense of, the vibe and not necessarily the service offered, to know if...I would feel comfortable or okay with it. Yeah. I think that's a big one, because it's definitely hard to review any kind of mental health services, especially because people go in in such different places with such different experiences...someone could have a horrible experience just because the person, the professional they were working with or the clinician just was not equipped to deal with that situation, but is amazing for someone else. So I think it's definitely kind of a balance there.* [P08]

Although participants valued others' input, many found it difficult to add resources and post reviews, so they did not use these crowdsourcing features. Participants gave a wide range of reasons they did not contribute. Some attributed their lack of engagement with these features to infrequent usage, lack of motivation, forgetfulness, or insufficient experience with mental health services:

> *I didn't because I didn't use it for that long...if I was using it for like a more consistent basis, then I would have been able to use it or potentially review any of the spots. Or maybe it's just—sometimes I forget. Honestly, I've not been one to review things a lot...sometimes I'd rather live in the moment than review it. So it could be a really unique aspect of it, if you do have users who are really consistent on reviewing things, but I don't think every user wants to review everything.* [P16]

The few participants who engaged in crowdsourcing appeared to do so because they were motivated to help fellow students.

When participants were asked to suggest ways to encourage engagement with crowdsourcing, a few acknowledged the

complexity and difficulty of motivating others. They were uncertain whether giving incentives or following up with service users would improve participation:

> *I don't know what you could use as a motivator...It doesn't have to be anything of any actual value, but having some sort of—an appearance of a reward at the end of something tends to motivate people, so that might be something that would maybe help.* [P09]

Peer-contributed information about mental health services and resources appeared to be a driver for user engagement with the app because this kind of information is hard to find elsewhere. Students wanted to hear from their peers to help them evaluate whether a service was the right fit.

Users Often Used the App in Response to an Immediate Need for Mental Health Information

User engagement appeared to be driven primarily by reactive rather than proactive behavior. Most participants reported using Thought Spot as a tool to learn about mental health resources during times of need. Several participants opened the app only when they were experiencing anxiety, depression, or other symptoms of poor mental health. As one user explained:

> *Actually, just when I need to, when I have some problem or issue or my friend happen[s] to ask me so I just show it to him....* [P12]

Some users reported using the app infrequently because they were not experiencing mental health issues during the 6-month study period. However, these users indicated that they would be willing to rely on Thought Spot if problems arose, as one user described:

> *I didn't end up going to any of them more than once or twice, even the thought of just having it there, knowing that I could use it if I wanted to, provided a level of comfort that helped when, you know, there were things that would make you spiral or you were not thinking very clearly or very logically.* [P09]

Although Thought Spot has features that can be used daily, such as adding reviews, crowdsourcing resources, and mood tracking, some users said they seldom engaged with these features.

> *I didn't really use the mood tracker...although I'm just bad at tracking things in general, so I guess in that way I could've used it regularly, but other than that, nah...because in terms of finding resources, your search kind of stops as soon as you've found something that works for you.* [P02]

While several participants said they did not regularly use Thought Spot to search for and access mental health support, they identified it as an option that they could rely on if they ever needed help.

Discussion

Principal Findings

Factors That Affected Postsecondary Students' Engagement With a Mental Health App

This study is among the few that describe factors affecting postsecondary students' attitudes, behaviors, engagement, and user experience when using a mental health mobile app. Participants identified the comprehensiveness and relevance of app content, user experience, integration with other apps, peer support, and reactive (versus proactive) behavior as factors that affected engagement. To varying degrees, these factors appeared to influence students' use of the Thought Spot app for mental health help-seeking and their willingness to use it in the future.

How engaged users were on Thought Spot appeared to be affected by factors that can be grouped into 5 themes. First, positive experiences on the app were tied to whether it delivered mental health information that users found to be concise, inclusive, relevant to their needs, and that included meaningful details. Dissatisfaction with the content appeared to decrease users' willingness to engage with the app and to use it to support their mental health. A second theme related to engagement was user experience and integration with other apps. Despite technical issues and lack of integration, some participants were still willing to engage with Thought Spot if they believed they could benefit from it. Third, some participants were motivated to engage with Thought Spot to support friends and family members who needed mental health support. Accessing crowdsourced information was a fourth theme related to engagement. Some participants used the app to read reviews about their peers' experiences with mental health resources and services. Fifth, engagement with the app appeared to be driven by reactive rather than proactive behavior, that is, participants often used the app when they had an immediate need for mental health information. These themes provide important insights into factors that affect the engagement of postsecondary students with mobile apps in the context of seeking help for mental health issues.

Research Theories on Low Engagement With Mobile Mental Health Interventions

The findings of this study that relate to user engagement are consistent with those of previous research on mobile health and mobile mental health interventions. That research has proposed general theories for low engagement, and this study adds direct evidence to support several concepts related to engagement [8,15,18,19]. For example, studies have theorized that usability issues, not being user-centered, and lacking relevant information about mental health services limited users' ability to address mental health problems or progress toward their help-seeking and wellness goals [18,19]. The themes developed in this study that relate to Thought Spot's content and user experience provide some evidence that supports existing theories of low engagement. For example, several users said that experiencing usability issues such as technical glitches or lacking integration hurt their subjective experience with using Thought Spot, which could thereby jeopardize engagement. Similarly, having

difficulties navigating the app and finding relevant content could have impaired engagement because several participants reported it as a source of user dissatisfaction.

Information Exchange With Peers and Family Members

The findings of this study complement existing research about barriers to and facilitators of mental health help-seeking among youth, specifically the importance of peer and family relationships [35-37]. For example, previous research has found that many young people prefer informal sources of support, often family and friends, when they are seeking help or information about mental health [35-37]. Participants in this study also expressed that preference and indicated that they would use the app to help family or friends access mental health information and services. Some participants described Thought Spot as a source of accessible and accurate information about mental health services and wanted improved ways to share this knowledge with others. Although sharing and communication features are not part of the current app version, participants in the co-design and prototyping phase of the project recommended embedding peer-to-peer communication within the app.

Although sharing and communication features were not available, participants still found a way to exchange mental health information with family and peers. This type of peer-to-peer communication could be seen as an innovative way of working around the app's limited features. Sharing content from Thought Spot suggested the potential of apps to improve awareness of mental health services and resources, which is the most common knowledge-related barrier to seeking help [35]. Moreover, the findings reiterate the crucial role that peers and family members often play during help-seeking [35-37].

The findings are also consistent with theories that identify peer support as a way to improve user engagement [19]. Several participants described crowdsourcing as an appealing strategy for students. Peer reviews of mental health services were a valuable decision-making aid for some participants who were seeking help for themselves. They explained that these kinds of reviews are scarce, but that they are more relevant and trustworthy than information from an unknown source. The request for more reviews from mental health service users underscores the importance of peer support as a feature that increases user engagement. Overall, the study provides new evidence to support existing theories about low engagement with mobile health and mobile mental health interventions.

Challenges of Co-Designing Apps

The Thought Spot project was student led and many postsecondary students were actively involved in deciding what features to include in the app that was evaluated during the RCT. Several findings from the participant interviews echo what students had discussed during the co-design workshops and focus groups in the earlier stages of Thought Spot's development [25]. At that time, students suggested adding more peer support, including features that would enable communication between users and with social groups. They also requested more information about service costs, accessibility, and languages spoken [25]. The research team considered these suggestions for the optimization phase of the app's development, but they could not be implemented for various reasons. For example, the databases from which Thought Spot draws information do not collect information on cost of services, accessibility, or wait times. Project cost constraints made it unfeasible to add complex features such as integration with other apps. Direct peer-to-peer interaction features were not implemented because they pose a high risk for misuse and the research team lacked the resources to monitor this activity for safety. Nonetheless, the similarities between students' perspectives during the development of Thought Spot and after the RCT show that participatory co-design research can be a useful tool for identifying key features that influence user engagement in the final product.

This study is also one of the few studies of mobile mental health interventions that points to the challenges during the co-design process of balancing user needs and perspectives with project resources, feasibility, and risk [38,39]. In future studies, it may be valuable for other researchers to also discuss the consequences when co-design suggestions cannot be implemented. Likewise, it can be useful to learn about the complex decision-making process that developers undergo when choosing what features to include or exclude. Doing so could identify areas of caution and guide other mobile mental health app developers.

Comparison With Mobile Mental Health Assessment Frameworks

Mobile mental health assessment frameworks, such as those developed by Chan et al [40], Zelmer et al [41], and Stoyanov et al [42], help researchers evaluate apps and guide developers in building high-quality, safe, and effective tools. These frameworks describe key considerations, including fit to target group, functionality, information quality, integration, user-centeredness, usefulness, usability, security, and transparency [40-42]. Although the frameworks are useful guidelines, it is unclear how much engagement with the technology will change when the framework criteria are satisfied [40-42]. The findings of this study provide preliminary indications, given the similarities between several criteria in the 3 frameworks cited above and the themes developed in this study. For example, functionality and usability criteria, which refer to an app's performance, reliability, and ease of use, are similar to the themes that emerged in this study that relate to user experience and willingness to use Thought Spot [40-42]. Likewise, framework criteria about fit to target group, information quality, and usefulness are reflected in this study's theme that links engagement with the provision of app content that is detailed, inclusive, and relevant [40-42]. These complementary findings indicate that content and usability improve engagement, but further investigation is required to measure the impact.

Findings from this study suggest that some participants prioritize usefulness over the user experience. However, the assessment frameworks described above do not rank the importance of each criterion [40-42]. It may be useful for future studies to explore the relative impact that each criterion has on engagement. That knowledge could help app developers determine what features or functions to prioritize to maximize adoption and engagement. Moreover, incorporating this information into existing

assessment frameworks could increase their practical value in guiding the development of projects with limited resources and time constraints, such as Thought Spot and many publicly funded co-designed projects [39].

It is important to note that sustained usage is not guaranteed even if all framework criteria are satisfied. For example, in our study, some students used Thought Spot only during times of pressing need, which could result in infrequent and sparse engagement, regardless of the quality or usefulness of the app. This behavior suggests that user engagement is context dependent, and that an app can be useful despite low engagement with it.

Limitations

This study had several limitations. Because of purposeful sampling, the participants may not be representative of all users in the RCT. In addition, the study did not factor mental illness diagnoses into the recruitment strategy or thematic analysis, which means that we may not have captured the perspectives of students who are in greatest need of mental health support. During the recruitment process, the research team was not able to engage students who did not use the app, so our analysis did not include feedback from the most disengaged and disinterested students. Lastly, the findings may not be fully generalizable to mental health solutions with different functions. For example, Thought Spot functions primarily as a stand-alone app to assist with finding and navigating to mental health resources, but the factors that encourage engagement with it may differ from the factors that encourage engagement with an app that involves direct communication with a mental health professional (eg, cognitive behavioral therapy or counseling apps).

Conclusions

This study demonstrates that content, usability, user-centeredness, and peer-to-peer communication are determinants of engagement with apps such as Thought Spot among postsecondary students. Failing to meet participants' expectations on these dimensions led to disengagement with the app. The findings highlight the challenges of balancing user needs and perspectives with project resources, feasibility, and risk during the co-design process, as well as the difficulty in predicting which app features will be successful, even after a thorough co-design process. The findings of this study support criteria for engagement proposed in several mobile mental health assessment frameworks. However, neither this study nor existing theoretical frameworks have determined whether certain criteria have a greater impact on engagement than others. Future studies that measure the relative importance of each criterion for user engagement would yield insights that could help app developers prioritize certain features or functions, creating interventions with greater engagement and that reflect what students want and need when they are seeking mental health support.

Acknowledgments

This project was supported through funding from the eHealth Innovation Partnership Program (eHIPP) grant from the Canadian Institute for Health Research (EH1-143558). The Canadian Institute for Health Research was not involved in the study design, data analysis, data interpretation, or writing of the report. We thank all the students who participated in the end-of-study interviews, randomized trial, and engagement activities during the development and evaluation of Thought Spot. We also thank all CAMH staff, collaborators, CAMH REDCap Operations Committee, youth advisory committee members, and practicum students who were involved since the beginning of the project, including Daryl Boshart, Almerinda Collela, Paulysha De Gannes, Genevieve Ferguson, Olivia Heffernan, Tyson Herzog, Mandee Kaur, Julia Roy, Shehab Sennah, Sarah Sharpe, Angie Solomon, Chelsea Stunden, Vivetha Thambinathan, Tim Tripp, Nicole vanHeerwaarden, Julie Zasada, and Hema Zbogar. We also thank QoC Health, the software developers of the Thought Spot app. AV currently receives funding from the National Institute of Mental Health, Canadian Institutes of Health Research, Canada Foundation for Innovation, CAMH Foundation, and the University of Toronto.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide and probing questions.
[DOCX File , 16 KB - mental_v8i4e23447_app1.docx]

References

1. 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] [Medline: 27484622]
2. Anderson KK, Norman R, MacDougall A, Edwards J, Palaniyappan L, Lau C, et al. Effectiveness of Early Psychosis Intervention: Comparison of Service Users and Nonusers in Population-Based Health Administrative Data. Am J Psychiatry 2018 May 01;175(5):443-452. [doi: 10.1176/appi.ajp.2017.17050480] [Medline: 29495897]
3. Kessler RC, Avenevoli S, Ries Merikangas K. Mood disorders in children and adolescents: an epidemiologic perspective. Biological Psychiatry 2001 Jun;49(12):1002-1014. [doi: 10.1016/s0006-3223(01)01129-5]

4. Seko Y, Kidd S, Wiljer D, McKenzie K. Youth mental health interventions via mobile phones: a scoping review. Cyberpsychol Behav Soc Netw 2014 Sep;17(9):591-602. [doi: 10.1089/cyber.2014.0078] [Medline: 25007383]
5. Garrido S, Cheers D, Boydell K, Nguyen QV, Schubert E, Dunne L, et al. Young People's Response to Six Smartphone Apps for Anxiety and Depression: Focus Group Study. JMIR Ment Health 2019 Oct 02;6(10):e14385 [FREE Full text] [doi: 10.2196/14385] [Medline: 31579023]
6. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual Research Review: Digital health interventions for children and young people with mental health problems: a systematic and meta-review. J Child Psychol Psychiatry 2016 Dec 10. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
7. Harrer M, Adam SH, Baumeister H, Cuijpers P, Karyotaki E, Auerbach RP, et al. Internet interventions for mental health in university students: A systematic review and meta-analysis. Int J Methods Psychiatr Res 2019 Jun;28(2):e1759. [doi: 10.1002/mpr.1759] [Medline: 30585363]
8. Becker TD, Torous JB. Recent Developments in Digital Mental Health Interventions for College and University Students. Curr Treat Options Psych 2019 Jun 14;6(3):210-220. [doi: 10.1007/s40501-019-00178-8]
9. Conley CS, Durlak JA, Shapiro JB, Kirsch AC, Zahniser E. A Meta-Analysis of the Impact of Universal and Indicated Preventive Technology-Delivered Interventions for Higher Education Students. Prev Sci 2016 Dec;17(6):659-678. [doi: 10.1007/s11121-016-0662-3] [Medline: 27225631]
10. Sander L, Rausch L, Baumeister H. Effectiveness of Internet-Based Interventions for the Prevention of Mental Disorders: A Systematic Review and Meta-Analysis. JMIR Ment Health 2016;3(3):e38 [FREE Full text] [doi: 10.2196/mental.6061] [Medline: 27535468]
11. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
12. 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 2016 Dec 13. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
13. 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] [Medline: 31573916]
14. Couper MP, Alexander GL, Zhang N, Little RJA, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]
15. 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 Jun 05. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
16. Ng MM, Firth J, Minen M, Torous J. User Engagement in Mental Health Apps: A Review of Measurement, Reporting, and Validity. Psychiatr Serv 2019 Mar 27:appips201800519. [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
17. Dunbar MS, Sontag-Padilla L, Kase CA, Seelam R, Stein BD. Unmet Mental Health Treatment Need and Attitudes Toward Online Mental Health Services Among Community College Students. Psychiatr Serv 2018 May 01;69(5):597-600. [doi: 10.1176/appi.ps.201700402] [Medline: 29540117]
18. Vaghefi I, Tulu B. The Continued Use of Mobile Health Apps: Insights From a Longitudinal Study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
19. Torous JB, Chan SR, Gipson SYT, Kim JW, Nguyen T, Luo J, et al. A Hierarchical Framework for Evaluation and Informed Decision Making Regarding Smartphone Apps for Clinical Care. Psychiatr Serv 2018 May 01;69(5):498-500. [doi: 10.1176/appi.ps.201700423] [Medline: 29446337]
20. Wiljer D, Abi-Jaoude A, Johnson A, Ferguson G, Sanches M, Levinson A, et al. Enhancing Self-Efficacy for Help-Seeking Among Transition-Aged Youth in Postsecondary Settings With Mental Health and/or Substance Use Concerns, Using Crowd-Sourced Online and Mobile Technologies: The Thought Spot Protocol. JMIR Res Protoc 2016 Nov 04;5(4):e201 [FREE Full text] [doi: 10.2196/resprot.6446] [Medline: 27815232]
21. Hagen P, Collin P, Metcalf A, Nicholas M, Rahilly K, Swainston N. Participatory design of evidence-based online youth mental health promotion, intervention and treatment. 2012. URL: http://pandora.nla.gov.au/pan/141862/20160405-1343/www.youngandwellcrc.org.au/wp-content/uploads/2014/03/Young_and_Well_CRC_IM_PD_Guide.pdf [accessed 2021-03-06]
22. Bandura A. Social cognitive theory of personality. In: Pervin LA, John OP, editors. Handbook of Personality (2nd ed). New York, NY: The Guilford Press; 1999:154-196.
23. Rickwood D, Thomas K. Conceptual measurement framework for help-seeking for mental health problems. Psychol Res Behav Manag 2012;5:173-183 [FREE Full text] [doi: 10.2147/PRBM.S38707] [Medline: 23248576]
24. Sennah S, Shi J, Hollenberg E, Johnson A, Ferguson G, Abi-Jaoudé A, et al. Thought Spot: Embedding Usability Testing into the Development Cycle. Stud Health Technol Inform 2019;257:375-381. [Medline: 30741226]

25. VanHeerwaarden N, Ferguson G, Abi-Jaoude A, Johnson A, Hollenberg E, Chaim G, et al. The Optimization of an eHealth Solution (Thought Spot) with Transition-Aged Youth in Postsecondary Settings: Participatory Design Research. J Med Internet Res 2018 Mar 06;20(3):e79 [FREE Full text] [doi: 10.2196/jmir.8102] [Medline: 29510970]
26. Wiljer D, Johnson A, McDiarmid E, Abi-Jaoude A, Ferguson G, Hollenberg E, et al. Thought Spot: Co-Creating Mental Health Solutions with Post-Secondary Students. Stud Health Technol Inform 2017;234:370-375. [Medline: 28186070]
27. Wiljer D, Shi J, Lo B, Sanches M, Hollenberg E, Johnson A, et al. Effects of a Mobile and Web App (Thought Spot) on Mental Health Help-Seeking Among College and University Students: Randomized Controlled Trial. J Med Internet Res 2020 Oct 30;22(10):e20790 [FREE Full text] [doi: 10.2196/20790] [Medline: 33124984]
28. Kim H, Sefcik JS, Bradway C. Characteristics of Qualitative Descriptive Studies: A Systematic Review. Res Nurs Health 2017 Feb;40(1):23-42 [FREE Full text] [doi: 10.1002/nur.21768] [Medline: 27686751]
29. Gao M, Kortum P, Oswald F. Psychometric Evaluation of the USE (Usefulness, Satisfaction, and Ease of use) Questionnaire for Reliability and Validity. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2018 Sep 27;62(1):1414-1418. [doi: 10.1177/1541931218621322]
30. Smith J, Harré R, Van LL. Rethinking Methods in Psychology. Thousand Oaks, CA: Sage; 1995.
31. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
32. Brooks J, McCluskey S, Turley E, King N. The Utility of Template Analysis in Qualitative Psychology Research. Qual Res Psychol 2015 Apr 03;12(2):202-222 [FREE Full text] [doi: 10.1080/14780887.2014.955224] [Medline: 27499705]
33. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
34. Assarroudi A, Heshmati Nabavi F, Armat MR, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. Journal of Research in Nursing 2018 Jan 10;23(1):42-55. [doi: 10.1177/1744987117741667]
35. Brown A, Rice SM, Rickwood DJ, Parker AG. Systematic review of barriers and facilitators to accessing and engaging with mental health care among at-risk young people. Asia Pac Psychiatry 2016 Mar;8(1):3-22. [doi: 10.1111/appy.12199] [Medline: 26238088]
36. Findlay LC, Sunderland A. Professional and informal mental health support reported by Canadians aged 15 to 24. Health Rep 2014 Dec;25(12):3-11 [FREE Full text] [Medline: 25517935]
37. Reavley NJ, Yap MBH, Wright A, Jorm AF. Actions taken by young people to deal with mental disorders: findings from an Australian national survey of youth. Early Interv Psychiatry 2011 Nov;5(4):335-342. [doi: 10.1111/j.1751-7893.2011.00292.x] [Medline: 21951823]
38. 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] [Medline: 29362208]
39. 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/fpsyt.2018.00481] [Medline: 30405450]
40. Chan S, Torous J, Hinton L, Yellowlees P. Towards a Framework for Evaluating Mobile Mental Health Apps. Telemed J E Health 2015 Jul 14. [doi: 10.1089/tmj.2015.0002] [Medline: 26171663]
41. Zelmer J, van HK, Notarianni M, van MT, Schellenberg M, Tannenbaum C. An Assessment Framework for e-Mental Health Apps in Canada: Results of a Modified Delphi Process. JMIR Mhealth Uhealth 2018 Jul 09;6(7):e10016 [FREE Full text] [doi: 10.2196/10016] [Medline: 29986846]
42. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]

Abbreviations

RCT: randomized controlled trial
USE: Usefulness, Satisfaction, and Ease of Use

Edited by G Eysenbach; submitted 18.08.20; peer-reviewed by N Wang, M Yuhas; comments to author 15.09.20; revised version received 12.11.20; accepted 07.12.20; published 02.04.21.

Please cite as:
Wong HW, Lo B, Shi J, Hollenberg E, Abi-Jaoude A, Johnson A, Chaim G, Cleverley K, Henderson J, Levinson A, Robb J, Voineskos A, Wiljer D
Postsecondary Student Engagement With a Mental Health App and Online Platform (Thought Spot): Qualitative Study of User Experience
JMIR Ment Health 2021;8(4):e23447
URL: https://mental.jmir.org/2021/4/e23447
doi:10.2196/23447
PMID:33797395

Original Paper

Virtual Reality–Based Psychotherapy in Social Anxiety Disorder: fMRI Study Using a Self-Referential Task

Ji-Won Hur[1], PhD; Hyemin Shin[1], BA; Dooyoung Jung[2], MD, PhD; Heon-Jeong Lee[3], MD, PhD; Sungkil Lee[4], PhD; Gerard J Kim[5], PhD; Chung-Yean Cho[6], MFA; Seungmoon Choi[7], PhD; Seung-Moo Lee[6], MFA; Chul-Hyun Cho[8,9], MD, PhD

[1]School of Psychology, Korea University, Seoul, Republic of Korea

[2]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea

[3]Department of Psychiatry, Korea University College of Medicine, Seoul, Republic of Korea

[4]Department of Software, Sungkyunkwan University, Suwon, Republic of Korea

[5]Digital Experience Laboratory, Department of Computer Science and Engineering, Korea University, Seoul, Republic of Korea

[6]Department of Film & Multimedia, Korea National University of Arts, Seoul, Republic of Korea

[7]Department of Computer Science and Engineering, Pohang University of Science and Technology, Pohang, Republic of Korea

[8]Department of Psychiatry, College of Medicine, Chungnam National University, Daejeon, Republic of Korea

[9]Department of Psychiatry, Chungnam National University Sejong Hospital, Sejong, Republic of Korea

Corresponding Author:
Chul-Hyun Cho, MD, PhD
Department of Psychiatry
College of Medicine
Chungnam National University
82 Munhwa-ro, Jung-gu
Daejeon, 35015
Republic of Korea
Phone: 82 44 995 4775
Email: david0203@gmail.com

Abstract

Background: Although it has been well demonstrated that the efficacy of virtual reality therapy for social anxiety disorder is comparable to that of traditional cognitive behavioral therapy, little is known about the effect of virtual reality on pathological self-referential processes in individuals with social anxiety disorder.

Objective: We aimed to determine changes in self-referential processing and their neural mechanisms following virtual reality treatment.

Methods: We recruited participants with and without a primary diagnosis of social anxiety disorder to undergo clinical assessments (Social Phobia Scale and Post-Event Rumination Scale) and functional magnetic resonance imaging (fMRI) scans. Participants with social anxiety disorder received virtual reality–based exposure treatment for 6 sessions starting immediately after baseline testing. After the sixth session, participants with social anxiety disorder completed follow-up scans during which they were asked to judge whether a series of words (positive, negative, neutral) was relevant to them.

Results: Of 25 individuals with social anxiety disorder who participated in the study, 21 completed the sessions and follow-up; 22 control individuals also participated. There were no significant differences in age (P=.36), sex (P=.71), or handedness (P=.51) between the groups. Whole-brain analysis revealed that participants in the social anxiety disorder group had increased neural responses during positive self-referential processing in the medial temporal and frontal cortexes compared with those in the control group. Participants in the social anxiety disorder group also showed increased left insular activation and decreased right middle frontal gyrus activation during negative self-referential processing. After undergoing virtual reality–based therapy, overall symptoms of the participants with social anxiety disorder were reduced, and these participants exhibited greater activity in a brain regions responsible for self-referential and autobiographical memory processes while viewing positive words during postintervention fMRI scans. Interestingly, the greater the blood oxygen level dependent changes related to positive self-referential processing, the lower the tendency to ruminate on the negative events and the lower the social anxiety following the virtual reality session.

Compared with that at baseline, higher activation was also found within broad somatosensory areas in individuals with social anxiety disorder during negative self-referential processing following virtual reality therapy.

Conclusions: These fMRI findings might reflect the enhanced physiological and cognitive processing in individuals with social anxiety disorder in response to self-referential information. They also provide neural evidence of the effect of virtual reality exposure therapy on social anxiety and self-derogation.

(JMIR Ment Health 2021;8(4):e25731) doi:10.2196/25731

KEYWORDS

virtual reality; VR; social anxiety; social phobia; exposure therapy; fMRI; unctional magnetic resonance imaging

Introduction

Social anxiety disorder (also known as social phobia) is characterized by a persistent fear of social situations in which the person would be exposed to possible scrutiny by others [1], with a lifetime prevalence of 2% to 7% in adults and a prevalence up to 25% in university students [2]. Individuals who have social anxiety disorder are typically hypervigilant for real or imagined feedback and more likely to rely on safety behaviors, aiming to avoid any stimuli or events that could trigger their social anxiety [3,4].

An elaborate cognitive–behavioral model proposed that pathological self-referential processing is an important factor in the developing and continuing to experience social anxiety disorder [5]. Increased activity in cortical midline structures (eg, medial prefrontal cortex, posterior cingulate cortex) and limbic areas (eg, amygdala, anterior cingulate cortex, insula) has been linked to biased self-referential processing [5,6] in social anxiety disorder. Altered activation in these brain areas is also correlated with abnormal self-focused attention [7], which includes socially anxious individuals' fear of being evaluated [8].

Interestingly, individuals with social anxiety disorder seem to be vulnerable to any type of evaluation toward them. That is, they are sensitive to potential negative evaluations from others (fear of negative evaluation), but also, positive evaluations from others [9] given that a positive social reputation inevitably elevates the individual's social status, which brings more scrutiny. Empirical evidence indicates that receiving positive social feedback is a horrific experience for patients with social anxiety disorder [10,11]. In addition, it has been shown that the higher the fear of positive evaluation, the higher the social anxiety and discomfort and the lower the assertiveness and perceived accuracy of the feedback [10,12]. Findings on the fear of evaluation in social anxiety disorder have been paralleled by a growing interest in their neurophysiological mechanisms. Previous studies [12] have suggested that aberrant neural activity may exists when individuals with social anxiety disorder receive feedback. Research using positive referential processing has also shown higher activation in bilateral medial prefrontal and inferior frontal cortices, fusiform gyrus, thalamus, left posterior superior temporal gyrus [13], and left posterior insula [14] in social anxiety disorder.

Unfortunately, social anxiety disorder affects self-concept and individual performance directly and long-term [15,16]. Therefore, timely diagnosis and intervention are necessary for those affected by social anxiety disorder. A large body of literature has shown that psychoeducation, imaginal and in vivo exposure, and assertive training can reduce social anxiety symptoms and hyperactivation in brain areas in social anxiety disorder [17-19]; however, a majority of individuals do not place social phobia as a priority for treatment, as they become distracted by other coexisting diseases or regard their symptoms as an inherited temperament such as shyness [20].

In recent years, virtual reality (VR) therapeutics, developed to overcome restricted accessibility to conventional psychotherapies, have been widely used as a valid and effective platform for patients with social anxiety disorder to learn evidence-based coping skills [21,22]. Meta-analyses [23,24] have shown a large effect size of VR exposure therapy for social anxiety disorder and performance anxiety; however, the extent of the effectiveness of social anxiety disorder–specific VR therapies on self-referential processing remains an open research question. Moreover, neurobiological evidence regarding VR-based interventions for social anxiety disorder is still lacking.

In this study, we aimed to assess positive and negative self-referential processing in individuals diagnosed with social anxiety disorder and to explore the effects of VR therapy on neural substrates related to self-referential processing. First, we hypothesized that alterations in self-referencing occur in social anxiety disorder. Second, we speculated that neuronal changes in certain brain regions occur during self-referential processing in individuals with social anxiety disorder who have undergone VR therapy. To assess this, we used VR therapy [25], an alternative to exposure therapy for individuals with social anxiety disorder that provides psychoeducational intervention tailored to social anxiety symptoms. In short, we expected to provide neural evidence of the efficacy and applicability of a VR-based therapeutic approach to alter self-referential processing in individuals with social anxiety disorder.

Methods

Recruitment

We recruited individuals with and without social anxiety disorder via advertisements posted online (eg, forums, social media, and a website). Participants with social anxiety disorder were eligible if they met the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria for social anxiety disorder, which was assessed with the Mini-International Neuropsychiatric Interview [26], and if they had a score ≥82

on the Korean version of the Social Avoidance and Distress Scale.

The exclusion criteria for all participants were (1) having a lifetime or current mental illness or neurological disorder that might elicit severe side effects from a VR experience (eg, schizophrenia spectrum disorder, bipolar disorder, posttraumatic stress disorder, panic disorder, substance use disorders, autism spectrum disorder, epilepsy, traumatic brain injury, suicide attempts); (2) having an intellectual disability (IQ <70; estimated with the short version of the Korean Wechsler Adult Intelligence Test Fourth Edition [27]); and (3) receiving psychotropic medication or psychotherapy at the time of research enrollment.

After they had been given a detailed explanation of the study, all participants provided written consent and completed MRI scanning safety eligibility screen.

VR-Based Psychotherapy

Composition and Contents of the Participatory VR Therapy Program

Individuals diagnosed with social anxiety disorder were asked to participate in the VR therapy program developed and verified by the authors [25]. Two psychiatrists and two licensed psychologists performed collaborative work to develop a social-anxiety scenario (during a team meeting in class), and the Art & Technology Lab at the Korean National University of Arts designed VR environments using the Unity game engine (Unity Technologies). Motion capture was conducted using Sensor Suite (Rokoko Electronics) for VR character animation, and voice acting was performed by 4 actors from Korean National University of Arts drama school. In our previous work [25], we proved that our participatory and interactive VR intervention had significant beneficial effects on depression, anxiety, shame, rumination, and social phobia.

The participatory VR intervention for social anxiety symptoms consisted of 3 stages (introduction, core, and finishing) and was divided into 3 levels of difficulty (easy, medium, and difficult). All participants used a VIVE (HTC Corporation) VR headset, and the participants' heart rates, skin galvanic response, and eye movements were measured during the VR experience.

In the introduction stage, the participants were invited to choose their avatar to learn how to use VR, calm their minds, and help them adapt to VR during the meditation-based warm-up. A voice guide was provided to help the participants relax and breathe while observing trees gently shaking to calm their minds. Through this, the participants not only adapted to the VR system but also moved to the next stage in a stable state of mind. The introduction stage was configured to take approximately 5 minutes.

The core stage was designed to provide a solution to individuals' exposure and participation in social anxiety situations by providing them with a VR environment in which a college student group was meeting for the first time to discuss an assigned task and introduce themselves. In the virtual setting, 7 to 8 nonplayer characters introduced themselves then the participants with social anxiety disorder were asked to introduce themselves. At the easy level, each participant took turns introducing themselves in a normal and calm manner, starting with the nonplayer characters. In the difficult situation of encountering an unfavorable reaction from nonplayer characters, the nonplayer characters became increasingly distracted while listening to the participant's self-introduction, making more small talk among themselves, concentrating harder on other tasks, or staring intently at each other. The medium level consisted of distractions between the easy and difficult levels. The core stage was configured to take approximately 7 to 8 minutes.

In the finishing stage, the participant, as in the introduction stage, once again experienced meditation-based VR that calmed the mind while observing a gently shaking tree and controlling their breathing. The VR program ended by providing general cognitive and behavioral psychoeducation for social anxiety disorder in both voice and text form via the VR system. The final stage was configured to take approximately 3 minutes.

Number of Participatory VR Solution Sessions and Rules

All participants were asked to complete a total of 6 VR sessions, each consisting of the 3 stages. In one visit, each participant could complete up to 2 VR sessions as long as the participant took at least a 2-hour break between sessions. All participants started at the easy VR level. From the second session onward, the participants were asked to select the level of difficulty they desired. The difficulty level was increased, maintained, or decreased at the participant's request to provide an individual-tailored intervention. The researchers stayed with the participants throughout the VR sessions to address emergencies such as extreme anxiety or panic attacks.

Experimental Procedure

Overview

All participants underwent fMRI while performing a self-referential processing task and completing self-report questionnaires, including the Korean version of the Social Phobia Scale (SPS) (H. Kim, unpublished) and Post-Event Rumination Scale (PERS) [28]. The participants with social anxiety disorder underwent fMRI and assessments before and after treatment, while the control participants underwent fMRI only once. The recommended sample size for a task fMRI is 20; our sample size (n=25 in the social anxiety disorder group at baseline, n=21 in the social anxiety disorder group at follow-up) seemed to have adequate statistical power [29]. The study was approved by the institutional review board of Korea University Anam Hospital in accordance with the Declaration of Helsinki. This study was registered (Clinical Research Information Service KCT0003854).

fMRI Experimental Task

A revised version of the Personal Relevance Rating Task (PRRT) was used for the fMRI task [30]. The task consisted of 2 runs with a duration of 9 minutes 22 seconds per run, and each run included 40 trials (yielding a total of 80 trials). Each stimulus word was projected onto an angled mirror mounted on the head coil for 2 seconds using E-prime software (Psychology Software Tools). Between experimental stimuli, a mask (row of X's; 10.8 seconds) and a fixation cue (1 second;

row of X's with prongs around the center X) were presented. The order of presentation of all stimuli was counterbalanced.

To build the set of stimuli, we selected 10 positive words, 10 negative words, and 20 neutral words for the experiment, balanced for arousal level, emotional valence, and word length. In addition to the normed emotional adjectives, 10 positive words and 10 negative words generated by the participants were added to the word list for the fMRI PRRT task. The instructions were as follows: “Generate ten positive words that best represent your strengths” and “Generate ten negative words that best represent your weaknesses.” All of these words in both normed stimuli (10 positive, 10 negative, 20 neutral words) and participant-generated stimuli (10 positive, 10 negative words) were selected from the Korean emotion words list [31].

During the fMRI, participants were instructed to make a judgment on each word as soon as possible by pressing the buttons for “not relevant to me,” “somewhat relevant to me,” or “relevant to me” after the word projected. Prior to the scan, participants completed a practice session to ensure that they understood the task. All responses and reaction times were recorded.

Image Acquisition and Analysis

Neuroimaging was performed using an MRI machine (3T Siemens Tim trio) equipped with a 12-channel head coil at the Brain Imaging Center, Korea University. Axial T2*-weighted images (echo time 26 ms, repetition time 2000 ms, flip angle 80°, field of view 210 mm, voxel size 2.5×2.5×3.4, slice thickness 3.4 mm, matrix size 84×84). A single functional run consisted of 278 volumes with 37 sequential axial slices each. In addition, structural T1-weighted images (208 slices; echo time 1.89 ms, repetition time 1670 ms, flip angle 9°, field of view 250 mm, matrix size 256×256) were also obtained to aid with spatial normalization.

Anatomic T1 and functional T2* images were analyzed with Statistical Parametric Mapping software (version 12; Welcome Department of Cognitive Neurology). The echo-planar images were corrected for slice acquisition time and realigned to correct for rigid body transformation, then the individual's anatomic image was coregistered to the mean functional image. The echo-planar images were subsequently normalized using nonlinear transformation parameters obtained by registering individual T1-weighted images to the Montreal Neurological Institute template [32] and smoothed with an isotropic 6-mm^3 full-width-at-half-maximum Gaussian kernel.

Statistical Analysis

We compared the baseline for demographic measures of each group using independent *t* tests (sex and handedness) or chi-square tests (age and education). Distributions for all clinical variables were analyzed for normality, skewness, and kurtosis prior to comparative analysis. We compared the baseline measurements of each group using parametric and nonparametric analysis as appropriate. Normally distributed variables (negative PERS score) were analyzed employing an analysis of variance (ANOVA), while variables found to have skewed distributions (SPS, positive PERS scores) were analyzed using Mann–Whitney *U* tests. We also conducted a multivariate analysis of variance (MANOVA) to compare behavioral data between groups. For within-group analysis, ANOVA was used for normally distributed SPS scores to identify differences between baseline and follow-up. Both positive and negative PERS scores before and after the intervention were compared using the Wilcoxon signed-rank test. A repeated measures ANOVA with time (baseline vs follow-up) and valence (positive, negative, neutral) as within-subject factors was used to evaluate changes in behavioral outcomes for participants with social anxiety disorder. To identify the brain regions responsible for positive and negative self-referential processing, first-level contrast images were created using the difference between the blood oxygen level-dependent (BOLD) signals recorded during each condition (positive-word images>neutral-word images; negative-word images>neutral-word images). At the second level, we conducted whole-brain analysis to assess which specific brain regions were sensitive to the conditions between the social anxiety disorder and control groups. Two-sample *t* tests were used to reveal group-related differences in brain activity between the social anxiety disorder and control groups. A paired *t* test was also used to estimate the main effect of the intervention for each condition between baseline and follow-up in the social anxiety disorder group. In the whole-brain analysis, significant activations were reported with an uncorrected threshold of $P<.001$ and a cluster extent threshold of $k\geq20$ voxels (equivalent to $t=3.55$), which is recommended to minimize the risk of type I (false-positive) errors [33,34]. This threshold is stricter than the uncorrected $P<.005$, $k\geq20$ voxel threshold, which is equivalent to a false discovery rate of .05 [35].

In social anxiety disorder group, the correlations between the percentage signal intensity changes following VR therapy sessions and clinical symptom scores were evaluated using the Spearman correlation coefficient. Statistical analysis of the demographic and clinical data was performed using SPSS Statistics (version 23.0; IBM Corp).

Results

Recruitment, Enrollment, and Completion

We initially recruited and enrolled 40 individuals with social anxiety disorder and 33 individuals for the control group who had no other neurological or psychiatric diagnoses; however, from the initial 73 individuals, 23 participants (social anxiety disorder: n=12; control: n=11) were excluded because of missing data (n=14), excessive head motion (n=6), or poor fMRI task performance due to loss of concentration, fatigue, or dizziness (n =3). Data from 4 additional participants were discarded due to head motion (n=1) and declined to be scanned (n=3) at the time of follow-up, and 3 participants with social anxiety disorder dropped out during the study prior completing all 6 sessions. Thus, the data from 25 individuals (15 women) with a primary diagnosis of social anxiety disorder and 22 controls (12 women) were used in the baseline analysis (Table 1), and there were no group differences in age, sex, or handedness. In addition, among the 25 individuals with social anxiety disorder, 21 participants completed the VR session and postintervention assessments.

Table 1. Sample characteristics and behavioral data.

Characteristic	Social anxiety disorder (n=25)	Controls (n=22)	Test statistic	*P* value
Sex, n			0.14 (1)[a]	.71
Male	15	12		
Female	10	10		
Age (years), mean (SD)	23.04 (3.35)	23.95 (3.50)	–0.92 (45)[b]	.36
Handedness, n			1.36 (2)[a]	.51
Left	0	1		
Right	23	20		
Both	2	1		
Education (years), mean (SD)	14.44 (1.45)	15.09 (1.88)	–1.34 (45)[b]	.19
Social Phobia Scale, mean (SD)	32.52 (13.17)	6.05 (5.26)	–5.67[c]	<.001
Negative PERS[d], mean (SD)	32.24 (11.18)	8.18 (6.23)	79.92 (1,45)[e]	<.001
Positive PERS[d], mean (SD)	13.84 (5.41)	26.64 (9.44)	4.28[c]	<.001
Personal Relevance Rating Task, mean (SD)				
Personal relevance rating				
Positive	1.97 (.57)	2.52 (.41)	14.23 (.24)[f]	<.001
Neutral	2.01 (.30)	1.80 (.35)	5.11 (.10)[f]	.03
Negative	2.24 (.42)	1.61 (.48)	22.53 (.33)[f]	<.001
Reaction time (milliseconds)				
Positive	1567.5 (842.7)	1174.8 (761.4)	2.78 (.06)[f]	.10
Neutral	1601.7 (784.9)	1270.2 (547.3)	2.75 (.06)[f]	.10
Negative	1539.8 (720.0)	1187.0 (549.0)	3.49 (.07)[f]	.07

[a]Chi square (*df*).

[b]*t* test statistic (*df*).

[c]Mann–Whitney *U* test *Z* statistic.

[d]PERS: Post-Event Rumination Scale. Task response ratings were scored on a 3-point Likert scale (1=not relevant; 2=somewhat relevant; 3=relevant).

[e]*F*(*df1*,*df2*).

[f]*F*(η^2_p).

Full Sample Description at Baseline

Demographic and Clinical Characteristics

Demographics did not differ significantly between the 2 groups (sex, *P*=.71; age, *P*=.36; handedness, *P*=.51; education, *P*=.19), but compared with the controls, individuals with social anxiety disorder had higher SPS (Mann–Whitney *U* test: *Z*=−5.67, *P*<.001, *r*=−.83) and negative PERS ($F_{1,47}$=79.92, *P*<.001, η^2_p=.64) scores as well as lower positive PERS scores (Mann–Whitney *U* test: *Z*=4.28, *P*<.001, *r*=−.63) (Table 1).

Behavioral Outcomes

Statistically significant multivariate effects were found between the social anxiety disorder and control groups for the response ratings and reaction times for words of every valence (positive, neutral, and negative) (Wilks lambda=.58, $F_{3,43}$=10.41, *P*<.001, η^2_p=.42). Participants who were highly socially anxious rated positive words as less ($F_{1,45}$=14.23, *P*<.001, η^2_p=.24) and negative words as more personally relevant ($F_{1,45}$=22.53, *P*<.001, η^2_p=.33) than controls did. There were no significant group differences in reaction time for any valence (Wilks lambda=.93, $F_{3,43}$=1.11, *P*=.36, η^2_p=.07) (Table 1).

Neural Correlates of Positive and Negative Self-Referential Processing at Baseline

In the whole-brain analysis, compared with the controls, the social anxiety disorder group showed increased activation of the bilateral inferior frontal gyrus and left middle temporal gyrus during positive self-referential processing (positive-word images>neutral-word images). In individuals with social anxiety disorder, there was also increased left insula activation and decreased right middle frontal gyrus activation in response to

negative self-referential processing (negative-word images>neutral-word images) (Figure 1 and Table 2).

Figure 1. (A) Statistical parametric map showing activation differences between the social anxiety disorder and control groups during positive self-referential processing. (B) Statistical parametric map showing activations associated with negative self-referential processing. HC: healthy control; IFG: inferior frontal gyrus; MFG: middle frontal gyrus; MTG: middle temporal gyrus; Neg: negative words; Neu: neutral words; Pos: positive words; SAD: social anxiety disorder. The color bar depicts Z values.

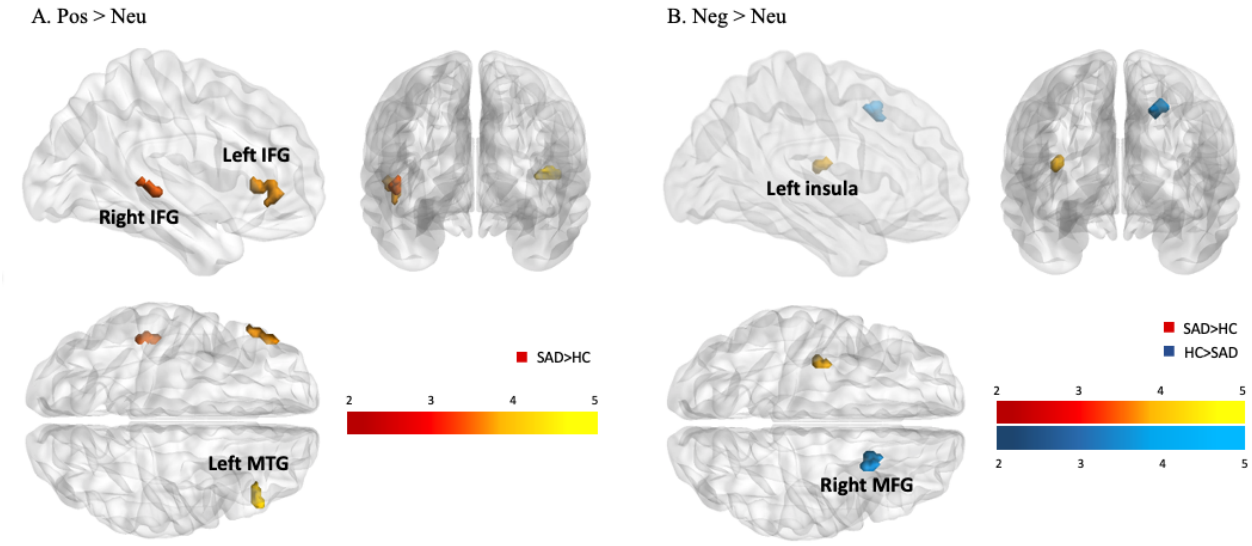


Table 2. Brain regions showing group differences in response to self-referential processing.

Valence and region	Coordinate			Statistic[a]		Social anxiety disorder with respect to control
	x	*y*	*z*	*k*	Z_{max}	
Positive>neutral words						
Inferior frontal gyrus						
Right pars triangularis	48	34	8	53	4.09	Greater activation
Left pars triangularis	–48	38	4	49	3.80	Greater activation
Left middle temporal gyrus	–48	–28	0	24	3.52	Greater activation
Negative>neutral words						
Left insula	–36	–16	14	23	3.66	Greater activation
Right middle frontal gyrus	22	14	46	41	4.05	Less activation

[a]$t \geq 3.28$, df=[1.0, 45.0].

VR-Based Treatment Responses in Social Anxiety Disorder

Behavioral Outcomes

Repeated measures ANOVA for the response ratings and reaction times in the social anxiety disorder group (n=21) showed marginally significant changes in the ratings for the positive words ($F_{1,19}$=3.85, P=.06, η^2_p=.16) and negative words ($F_{1,19}$=3.77, P=.07, η^2_p=.16). No statistically significant changes were found for reaction times before and after VR treatment (Table 3).

Table 3. Changes in clinical symptoms and behavior performance of participants with social anxiety disorder following VR therapy.

	Baseline (n=21)	Follow-up (n=21)	Test statistic	*P* value
Social Phobia Scale, mean (SD)	30.86 (13.34)	23.52 (12.34)	9.83 (1,20)[a]	.005
Negative PERS[b], mean (SD)	31.95 (11.66)	21.19 (11.62)	–3.32[c]	<.001
Positive PERS[b], mean (SD)	13.86 (5.12)	14.62 (7.61)	.32[c]	.75
Personal Relevance Rating Task, mean (SD)				
Personal relevance rating				
Positive	1.93 (.60)	2.17 (.61)	3.85 (.16)[d]	.06
Neutral	1.98 (.28)	2.05 (.30)	1.18 (.06)[d]	.29
Negative	2.18 (.41)	2.01 (.42)	3.77 (.16)[d]	.07
Reaction time (milliseconds)				
Positive	1594.0 (894.8)	1333.7 (458.3)	2.10 (.10)[d]	.16
Neutral	1677.0 (823.0)	1469.1 (661.0)	1.15 (.05)[d]	.30
Negative	1599.7 (743.0)	1369.7 (577.9)	1.99 (.09)[d]	.17

[a]*F*(*df1*,*df2*).

[b]PERS: Post-Event Rumination Scale. Task response ratings were scored on a 3-point (1 to 3) Likert scale.

[c]Wilcoxon signed-rank test *Z* statistic.

[d]$F(\eta^2_p)$.

Changes in Clinical Symptom Severity in Social Anxiety Disorder

The SPS and negative PERS scores of the social anxiety disorder group decreased remarkably ($F_{1,20}$=9.83, *P*=.005, η^2_p=.33; Wilcoxon signed-rank test: *Z*=–3.32, *P*<.001, *r*=.51, respectively) following VR therapy. There were no significant changes in positive PERS scores between baseline and follow-up (*P*=.75) (Table 3).

Changes in Neural Response to Self-Referential Processing in Social Anxiety Disorder

The social anxiety disorder group had significantly increased activation of positive self-referential stimuli (positive>neutral) in the right posterior cingulate cortex/precuneus, lingual gyrus, left inferior temporal gyrus, precentral gyrus, and postcentral gyrus after treatment. Moreover, increased activation was found during negative self-referential processing (negative>neutral) in the left middle occipital gyrus, parahippocampus, left Rolandic operculum, left superior frontal gyrus, and left caudate nucleus (Figure 2, Table 4) at follow-up compared with that at baseline.

Figure 2. Changes in neural responses during (A) positive self-referential processing and (B) negative self-referential processing among individuals with social anxiety disorder. ITG: inferior temporal gyrus; LG: lingual gyrus; MOG: middle occipital gyrus; Neg: negative words; Neu: neutral words; Pos: positive words; PCC: posterior cingulate cortex; RO: Rolandic operculum; SFG: superior frontal gyrus. The color bar depicts Z values.

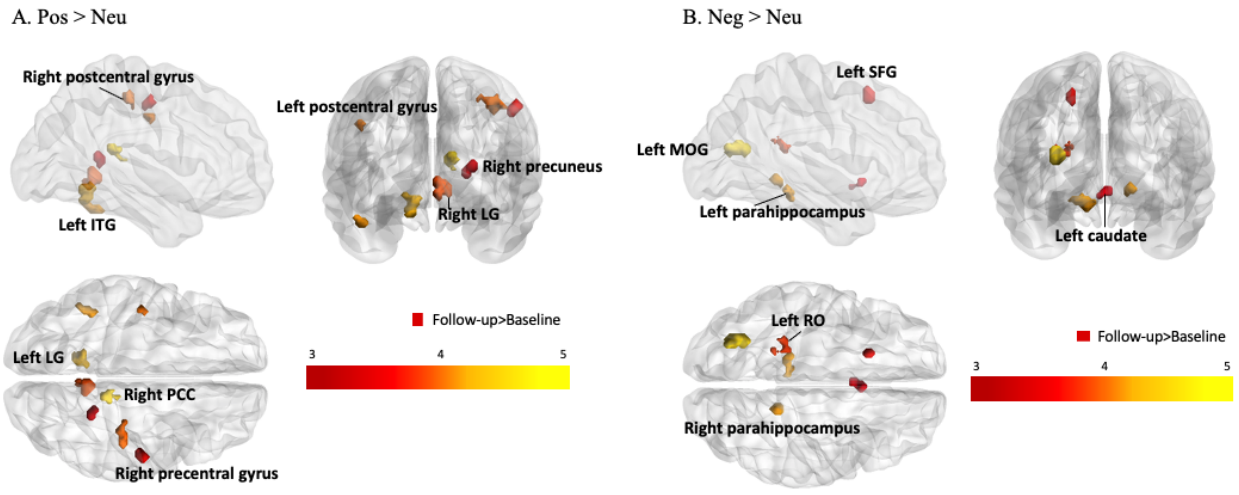


Table 4. Brain regions exhibiting significant changes in self-referential processing in social anxiety disorder following VR therapy.

Valence and region	Coordinate			Statistic[a]		Follow-up with respect to baseline
	X	*Y*	*Z*	*k*	Z_{max}	
Positive>neutral words						
Right posterior cingulate cortex/precuneus	12	–38	18	48	4.38	Greater activation
Right posterior cingulate cortex/precuneus	26	–48	12	22	3.45	Greater activation
Lingual gyrus						
Left lingual gyrus	–12	–54	–10	57	4.27	Greater activation
Right lingual gyrus	6	–50	–2	64	3.96	Greater activation
Left inferior temporal gyrus	–46	–50	–24	21	4.13	Greater activation
Postcentral gyrus						
Left postcentral gyrus	–44	–12	38	20	4.03	Greater activation
Right postcentral gyrus	34	–28	50	69	4.02	Greater activation
Right precentral gyrus	52	–14	48	25	3.81	Greater activation
Negative>neutral words						
Left middle occipital gyrus	–30	–64	16	77	4.40	Greater activation
Parahippocampus						
Left parahippocampus	–18	–36	–12	61	4.17	Greater activation
Right parahippocampus	16	–42	–4	21	4.10	Greater activation
Left Rolandic operculum	–24	–36	18	45	3.95	Greater activation
Left superior frontal gyrus	–20	18	52	23	3.88	Greater activation
Left caudate nucleus	0	12	–8	37	3.84	Greater activation

[a]$t \geq 3.55$, df=[1.0, 20.0].

Correlation Analyses Between Neural Changes During Self-Referential Processing and Clinical Measures in Social Anxiety Disorder Following VR Therapy

Changes in lingual gyrus activation associated with positive self-referential processing were related to alleviated social anxiety (SPS, Spearman ρ=–.52, *P*=.02) and decreased rumination on the negative events (negative PERS, Spearman ρ=–.61, *P*=.005) in the postintervention sessions. In addition, mean percentage BOLD signal changes in regions revealing significant changes following the VR intervention were not correlated with either positive PERS score at follow-up or the other symptom change rates in social anxiety disorder.

Discussion

General

Over the past decade, great endeavors have been made to provide VR-based therapeutic interventions for social anxiety disorder [36]. Individuals with social anxiety disorder continuously allocate their attentional resources to their self-evaluation and self-referential processing. Such cognitive processes can interfere with accurate perception and interpretation of the self and the social environment [37]. These individuals' excessive self-focused attention will eventually become impediments to social and professional achievements [38]. We aimed to demonstrate the neural correlates of self-referential processing in social anxiety disorder and the changes in brain activations following VR-based therapy, one of the promising interventions for social anxiety. The results demonstrated that individuals diagnosed with social anxiety disorder showed increased neural activity in response to both positive and negative self-referential stimuli. We also found that VR therapy alleviated anxiety symptoms and enhanced neural activity across a wide range of brain areas, including the frontal, temporal, and occipital regions, of individuals with social anxiety disorder. In particular, the SPS scores decreased significantly below the range of severe social anxiety (from mean 30.86, SD 13.34 at baseline to mean 23.52, SD 12.34 at follow-up; the clinical cutoff score for severe social anxiety is 24 [39]). In the correlation analyses, those who showed greater functional changes in the lingual gyrus during positive self-referential processing were observed to have less social anxiety and less engagement in negative rumination in the postintervention sessions. Our results might provide several insights regarding self-referential processing in social anxiety disorder in light of the legacy accumulated by brain imaging studies.

Increased Brain Activity During Self-Referential Processing

Group differences in neural areas recruited during the presentation of emotional self-referential stimuli were found. Many more areas, including the bilateral inferior frontal gyrus

and left middle temporal gyrus, were recruited during positive word processing (positive>neutral) than during negative self-referential word processing in social anxiety disorder. An increasing number of studies have proposed the fear of positive evaluation in social anxiety in addition to the fear of negative evaluation [10,40]. The increased inferior frontal gyrus and middle temporal gyrus activation reported herein further support previous findings that both regions are responsible for the down-regulation of socially driven emotions [41]. In particular, the left inferior frontal gyrus is well-known to be related to high selection demand among competing alternatives [42] as well as to self-referential processing [43]. This may indicate the possibility that the processing of positive self-referential information is a demanding task for socially anxious individuals.

Additionally, as hypothesized, negative self-referential processing (negative>neutral) was associated with stronger activation of the left insula in social anxiety disorder. This hyperactivation seems to reflect the aversive response or hypervigilance to negative self-referential stimuli [44-46].

VR Therapy

Changes in Neural Responses to Positive Words

Another overarching aim of this fMRI study was to identify neural changes associated with self-referential processing as well as anxiety symptom reduction in social anxiety disorder following VR therapy designed to teach self-assertiveness. Because the control group did not participate in the therapy sessions, it cannot be decisively stated that the fMRI data obtained were the direct results of the VR therapy intervention. What is clear, however, is that there were significant changes in brain activity during positive self-referencing processing in the social anxiety disorder group who completed the VR sessions. That is, the participants showed not only significant reductions in social anxiety disorder symptoms but also neural changes in the right posterior cingulate cortex/precuneus, lingual gyrus, left inferior temporal gyrus, postcentral gyrus, and right precentral gyrus during positive self-referential processing after VR therapy.

We found increased activity in the posterior cingulate cortex/precuneus of individuals with social anxiety disorder. These cortical regions, along with the precentral gyrus, exhibit enhanced activation during the processing of positive self-referential stimuli in healthy volunteers [30]. The posterior cingulate cortex/precuneus also plays a crucial role in integrating autobiographical information regarding the self [47]. Interestingly, a recent finding from experimental research suggested that social anxiety relates to impaired memory for social scenarios that ended with positive outcomes [48]. Altogether, our results may reflect increased cognitive efforts to access autobiographical memory while processing positive self-referential information in social anxiety disorder. This result needs to be further studied using an autobiographical memory task and comparison with a control group.

After VR therapy, the activation of the lingual gyrus and inferior temporal gyrus during positive self-referential processing also increased. According to previous studies [49,50] using healthy populations, stronger activations of the lingual gyrus, inferior temporal gyrus, and posterior cingulate cortex are closely related to self-referential processing. Intriguingly, the greater the change in lingual gyrus activation while processing positive self-referential information, the lower the tendency to ruminate on negative life events and the lower the social anxiety levels following the VR intervention session. Additionally, the inferior temporal gyrus has been reported to play an important role, along with the posterior cingulate cortex, in judging whether a series of stimuli is self-related [51]. Therefore, increased BOLD signals in these cortical regions may reflect the facilitative effect of VR therapy on the process by which the individuals with social anxiety disorder accept the positive-valence words as being relevant to them.

The precentral and postcentral gyri also showed increased activation during positive self-referential processing in individuals with social anxiety disorder upon completing the VR sessions. A similar brain activation pattern was observed in individuals with anxiety disorders who used cognitive reappraisal [52]. We conjecture that activation in the postcentral gyrus represent the therapeutic effects of the social anxiety–focused VR program. Further research is needed to investigate whether the therapeutic effect of our VR program could be extended to individuals with other clinical conditions that exhibit reduced postcentral gyrus activity, such as those with a generalized anxiety disorder [53] or a history of childhood abuse [54].

Changes in Neural Responses to Negative Words

Participants with social anxiety disorder exhibited greater activation in cortical and subcortical regions, which are known to be involved in somatosensory integration, during the processing of negative self-referential stimuli after VR therapy than at baseline. Similar findings were reported from an fMRI study [19,55] showing the neural mechanisms of the cognitive reappraisal of negative self-beliefs in individuals diagnosed with social anxiety disorder; in particular, the participants with social anxiety disorder demonstrated greater cognitive and somatosensory brain responses while reappraising negative self-beliefs. The most recent study using healthy adults [56] showed that cognitive bias modification for interpretation resulted in significantly greater activations of the somatomotor and somatosensory areas and occipital lobe. Our results may suggest that the VR-based therapy facilitated the perceptual or sensory-motor processing of negative self-referential stimuli in individuals with social anxiety disorder who had been willing to avoid negative feedback cues. Further research using a control group is needed to explore whether such an enhancement of brain activation can be linked to an approach-oriented strategy with negative-valence stimuli in social anxiety disorder.

Limitations

Several limitations should be noted. First, the case-control study using sham was not properly performed. Therefore, there is a limitation in interpretation to determine the effect of VR therapy through this study. Second, the lack of a direct comparison with conventional face-to-face therapy should be considered when interpreting these findings. A randomized controlled trial with a larger sample size is necessary to confirm the benefits of this VR therapy intervention on self-referential processing in social

anxiety disorder. Finally, it should be noted that some our showed patterns similar to those of previous studies of depression [57]. In future research, a transdiagnostic approach would be a reasonable way to explore a variety of diagnostic criteria for other conditions.

Conclusions

The body of literature supporting the effect of VR-based therapy for individuals with social anxiety disorder has been growing. Our study showed inefficient neural activation in a wide range of cortical regions, suggesting an increased predisposition for excessive cognitive processing in response to self-referential stimuli in social anxiety disorder. Following successful treatment with a VR intervention, symptoms in individuals with social anxiety disorder were reduced, which was demonstrated to be related to brain activation changes. Enhanced activations were also exhibited across brain regions that engage in self-image construction, autobiographical memory processing, and sensory information integration in healthy adults [30,47,56], indicating that VR may modulate the neural mechanisms responsible for self-reference. To our knowledge, this is the first neuroimaging study to specify the changes in the psychophysiological responses to self-referential information in social anxiety disorder in response to VR therapy. We believe that our findings may contribute to a better understanding of the therapeutic effects of VR-based interventions, which could be included in the routine treatment of social anxiety disorder.

Acknowledgments

This work was supported by a National Research Foundation of Korea, Ministry of Science and ICT (grant number NRF-2020R1C1C1007463), the ICT & Future Planning for Convergent Research in the Development Program for Research &Development Convergence over Science and Technology Liberal Arts (grant number NRF-2017M3C1B6070978), the Brain Research Program of the National Research Foundation of Korea (grant number 2017M3C7A1048040), and the Ulsan National Institute of Science & Technology Korea research brand program (grant number 1.200040.01).

Conflicts of Interest

None declared.

References

1. Social anxiety disorder (social phobia). In: Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Arlington, VA: American Psychiatric Association; 2013:202-208.
2. Dos Santos SR, de Lima Osório F, Martin-Santos R, Zuardi A, Hallak J, Crippa J. Modulation of the endocannabinoid and oxytocinergic systems as a potential treatment approach for social anxiety disorder. CNS Drugs 2019 Oct;33(10):1031-1038 [FREE Full text] [doi: 10.1007/s40263-019-00669-5] [Medline: 31617149]
3. Blair K, Blair R. A cognitive neuroscience approach to generalized anxiety disorder and social phobia. Emot Rev 2012 Apr 27;4(2):133-138 [FREE Full text] [doi: 10.1177/1754073911430251]
4. Leahy R. Emotional schemas and resistance to change in anxiety disorders. Cogn Behav Pract 2007 Feb;14(1):36-45 [FREE Full text] [doi: 10.1016/j.cbpra.2006.08.001]
5. Yoon H, Seo E, Kim J, Choo I. Neural correlates of self-referential processing and their clinical implications in social anxiety disorder. Clin Psychopharmacol Neurosci 2019 Feb 28;17(1):12-24 [FREE Full text] [doi: 10.9758/cpn.2019.17.1.12] [Medline: 30690936]
6. Lin Y, Callahan C, Moser J. A mind full of self: Self-referential processing as a mechanism underlying the therapeutic effects of mindfulness training on internalizing disorders. Neurosci Biobehav Rev 2018 Sep;92:172-186 [FREE Full text] [doi: 10.1016/j.neubiorev.2018.06.007] [Medline: 29886175]
7. Boehme S, Miltner W, Straube T. Neural correlates of self-focused attention in social anxiety. Soc Cogn Affect Neurosci 2015 Jun;10(6):856-862 [FREE Full text] [doi: 10.1093/scan/nsu128] [Medline: 25326038]
8. Spurr J, Stopa L. Self-focused attention in social phobia and social anxiety. Clin Psychol Rev 2002 Sep;22(7):947-975 [FREE Full text] [doi: 10.1016/s0272-7358(02)00107-1]
9. Wallace S, Alden L. Social anxiety and standard setting following social success or failure. Cogn Ther Res 1995 Dec;19(6):613-631 [FREE Full text] [doi: 10.1007/bf02227857]
10. Weeks J, Heimberg R, Rodebaugh T. The Fear of Positive Evaluation Scale: assessing a proposed cognitive component of social anxiety. J Anxiety Disord 2008;22(1):44-55 [FREE Full text] [doi: 10.1016/j.janxdis.2007.08.002] [Medline: 17884328]
11. Gilbert P. Evolution and social anxiety: the role of attraction, social competition, and social hierarchies. Psychiatr Clin 2001 Dec;24(4):723-751 [FREE Full text] [doi: 10.1016/s0193-953x(05)70260-4]
12. Budnick C, Kowal M, Santuzzi A. Social anxiety and the ironic effects of positive interviewer feedback. Anxiety Stress Coping 2015;28(1):71-87 [FREE Full text] [doi: 10.1080/10615806.2014.919386] [Medline: 24773204]
13. Goldin P, Ramel W, Gross J. Mindfulness meditation training and self-referential processing in social anxiety disorder: behavioral and neural effects. J Cogn Psychother 2009;23(3):242-257 [FREE Full text] [doi: 10.1891/0889-8391.23.3.242]

14. Miedl S, Blechert J, Klackl J, Wiggert N, Reichenberger J, Derntl B, et al. Criticism hurts everybody, praise only some: common and specific neural responses to approving and disapproving social-evaluative videos. Neuroimage 2016 May 15;132:138-147 [FREE Full text] [doi: 10.1016/j.neuroimage.2016.02.027] [Medline: 26892859]
15. Brook C, Willoughby T. The social ties that bind: social anxiety and academic achievement across the university years. J Youth Adolesc 2015 May;44(5):1139-1152 [FREE Full text] [doi: 10.1007/s10964-015-0262-8] [Medline: 25691148]
16. Vilaplana-Pérez A, Pérez-Vigil A, Sidorchuk A, Brander G, Isomura K, Hesselmark E, et al. Much more than just shyness: the impact of social anxiety disorder on educational performance across the lifespan. Psychol Med 2020 Jan 07:1-9 [FREE Full text] [doi: 10.1017/s0033291719003908]
17. Percell L, Berwick P, Beigel A. The effects of assertive training on self-concept and anxiety. Arch Gen Psychiatry 1974 Oct;31(4):502-504 [FREE Full text] [doi: 10.1001/archpsyc.1974.01760160052011] [Medline: 4419579]
18. Yang L, Zhou X, Pu J, Liu L, Cuijpers P, Zhang Y, et al. Efficacy and acceptability of psychological interventions for social anxiety disorder in children and adolescents: a meta-analysis of randomized controlled trials. Eur Child Adolesc Psychiatry 2019 Jan;28(1):79-89 [FREE Full text] [doi: 10.1007/s00787-018-1189-x] [Medline: 30006672]
19. Goldin P, Manber T, Hakimi S, Canli T, Gross J. Neural bases of social anxiety disorder: emotional reactivity and cognitive regulation during social and physical threat. Arch Gen Psychiatry 2009 Feb;66(2):170-180 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2008.525] [Medline: 19188539]
20. Ruscio A, Brown T, Chiu W, Sareen J, Stein M, Kessler R. Social fears and social phobia in the USA: results from the National Comorbidity Survey Replication. Psychol Med 2007 Nov 02;38(1):15-28 [FREE Full text] [doi: 10.1017/s0033291707001699]
21. Bouchard S, Dumoulin S, Robillard G, Guitard T, Klinger E, Forget H, et al. Virtual reality compared with exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. Br J Psychiatry 2017 Apr;210(4):276-283 [FREE Full text] [doi: 10.1192/bjp.bp.116.184234] [Medline: 27979818]
22. Verma R, Verma K. Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder. Br J Psychiatry 2018 Oct;213(4):617-617 [FREE Full text] [doi: 10.1192/bjp.2018.184] [Medline: 30247117]
23. Carl E, Stein A, Levihn-Coon A, Pogue J, Rothbaum B, Emmelkamp P, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019 Jan;61:27-36 [FREE Full text] [doi: 10.1016/j.janxdis.2018.08.003] [Medline: 30287083]
24. Kampmann I, Emmelkamp P, Morina N. Meta-analysis of technology-assisted interventions for social anxiety disorder. J Anxiety Disord 2016 Aug;42:71-84 [FREE Full text] [doi: 10.1016/j.janxdis.2016.06.007] [Medline: 27376634]
25. Kim H, Lee S, Jung D, Hur J, Lee H, Lee S, et al. Effectiveness of a participatory and interactive virtual reality intervention in patients With social anxiety disorder: longitudinal questionnaire study. J Med Internet Res 2020 Oct 06;22(10):e23024 [FREE Full text] [doi: 10.2196/23024] [Medline: 33021481]
26. Yoo S, Kim Y, Noh J, Oh K, Kim C, NamKoong K. Validity of Korean version of the mini-international neuropsychiatric interview. Anxiety Mood 2006;2(1):50-55.
27. Choe A, Hwang S, Kim J, Park K, Chey J, Hong S. Validity of the K-WAIS-IV Short Forms. Korean J Clin Psychol 2014 May;33(2):413-428. [doi: 10.15842/kjcp.2014.33.2.011]
28. Im S, Kwon SM, Choi H. Korean J Clin Psychol 2007 Feb;26(1):39-56. [doi: 10.15842/kjcp.2007.26.1.003]
29. Thirion B, Pinel P, Mériaux S, Roche A, Dehaene S, Poline JB. Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. Neuroimage 2007 Mar;35(1):105-120 [FREE Full text] [doi: 10.1016/j.neuroimage.2006.11.054] [Medline: 17239619]
30. Wagner G, Koch K, Schachtzabel C, Peikert G, Schultz C, Reichenbach J, et al. Self-referential processing influences functional activation during cognitive control: an fMRI study. Soc Cogn Affect Neurosci 2013 Oct;8(7):828-837 [FREE Full text] [doi: 10.1093/scan/nss074] [Medline: 22798398]
31. Park IJ, Min KW. Making a list of Korean emotion terms and exploring dimensions underlying them. Korean J Soc Pers Psychol 2005;19(1):109-129.
32. Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput Assist Tomogr 1994;18(2):192-205. [Medline: 8126267]
33. Eklund A, Nichols TE, Knutsson H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci USA 2016 Aug 08;113(28):7900-7905 [FREE Full text] [doi: 10.1073/pnas.1602413113]
34. Roiser J, Linden D, Gorno-Tempinin M, Moran R, Dickerson B, Grafton S. Minimum statistical standards for submissions to Neuroimage: Clinical. Neuroimage Clin 2016;12:1045-1047 [FREE Full text] [doi: 10.1016/j.nicl.2016.08.002] [Medline: 27995071]
35. Lieberman M, Cunningham W. Type I and Type II error concerns in fMRI research: re-balancing the scale. Soc Cogn Affect Neurosci 2009 Dec;4(4):423-428 [FREE Full text] [doi: 10.1093/scan/nsp052] [Medline: 20035017]
36. Emmelkamp P, Meyerbröker K, Morina N. Virtual reality therapy in social anxiety disorder. Curr Psychiatry Rep 2020 May 13;22(7):32-32 [FREE Full text] [doi: 10.1007/s11920-020-01156-1] [Medline: 32405657]
37. Bögels SM, Mansell W. Attention processes in the maintenance and treatment of social phobia: hypervigilance, avoidance and self-focused attention. Clin Psychol Rev 2004 Nov;24(7):827-856 [FREE Full text] [doi: 10.1016/j.cpr.2004.06.005] [Medline: 15501558]

38. Stein M, Stein D. Social anxiety disorder. Lancet 2008 Mar;371(9618):1115-1125 [FREE Full text] [doi: 10.1016/s0140-6736(08)60488-2]
39. Hope D, Heimberg R, Turk C. Managing Social Anxiety, Therapist Guide: A Cognitive-Behavioral Therapy Approach. New York: Oxford University Press; 2019.
40. Reichenberger J, Blechert J. Malaise with praise: a narrative review of 10 years of research on the concept of fear of positive evaluation in social anxiety. Depress Anxiety 2018 Dec;35(12):1228-1238 [FREE Full text] [doi: 10.1002/da.22808] [Medline: 30144225]
41. Grecucci A, Giorgetta C, Bonini N, Sanfey A. Reappraising social emotions: the role of inferior frontal gyrus, temporo-parietal junction and insula in interpersonal emotion regulation. Front Hum Neurosci 2013;7:523 [FREE Full text] [doi: 10.3389/fnhum.2013.00523] [Medline: 24027512]
42. Kan I, Thompson-Schill S. Effect of name agreement on prefrontal activity during overt and covert picture naming. Cogn Affect Behav Neurosci 2004 Mar;4(1):43-57 [FREE Full text] [doi: 10.3758/cabn.4.1.43] [Medline: 15259888]
43. Morin A, Michaud J. Self-awareness and the left inferior frontal gyrus: inner speech use during self-related processing. Brain Res Bull 2007 Nov 01;74(6):387-396 [FREE Full text] [doi: 10.1016/j.brainresbull.2007.06.013] [Medline: 17920447]
44. Yoon H, Kim J, Shin YB, Choi SH, Lee SK, Kim JJ. Neural activity during self-referential working memory and the underlying role of the amygdala in social anxiety disorder. Neurosci Lett 2016 Aug 03;627:139-147 [FREE Full text] [doi: 10.1016/j.neulet.2016.05.068] [Medline: 27260987]
45. Wong Q, Rapee R. The developmental psychopathology of social anxietyphobia in adolescents. In: Ranta K, La Greca A, Garcia-Lopez LJ, Marttunen M, editors. Social Anxiety and Phobia in Adolescents. Cham: Springer; 2015:11-37.
46. Spence S, Rapee R. The etiology of social anxiety disorder: an evidence-based model. Behav Res Ther 2016 Nov;86:50-67 [FREE Full text] [doi: 10.1016/j.brat.2016.06.007] [Medline: 27406470]
47. Rohde K, Caspar F, Koenig T, Pascual-Leone A, Stein M. Neurophysiological traces of interpersonal pain: how emotional autobiographical memories affect event-related potentials. Emotion 2018 Mar;18(2):290-303 [FREE Full text] [doi: 10.1037/emo0000356] [Medline: 28857583]
48. Romano M, Tran E, Moscovitch D. Social anxiety is associated with impaired memory for imagined social events with positive outcomes. Cogn Emot 2020 Jun;34(4):700-712 [FREE Full text] [doi: 10.1080/02699931.2019.1675596] [Medline: 31595830]
49. Komulainen E, Heikkilä R, Meskanen K, Raij T, Nummenmaa L, Lahti J, et al. A single dose of mirtazapine attenuates neural responses to self-referential processing. J Psychopharmacol 2016 Jan;30(1):23-32 [FREE Full text] [doi: 10.1177/0269881115616384] [Medline: 26577062]
50. Benoit R, Gilbert S, Volle E, Burgess P. When I think about me and simulate you: medial rostral prefrontal cortex and self-referential processes. Neuroimage 2010 Apr 15;50(3):1340-1349 [FREE Full text] [doi: 10.1016/j.neuroimage.2009.12.091] [Medline: 20045478]
51. Herold D, Spengler S, Sajonz B, Usnich T, Bermpohl F. Common and distinct networks for self-referential and social stimulus processing in the human brain. Brain Struct Funct 2016 Sep;221(7):3475-3485 [FREE Full text] [doi: 10.1007/s00429-015-1113-9] [Medline: 26365506]
52. Picó-Pérez M, Radua J, Steward T, Menchón J, Soriano-Mas C. Emotion regulation in mood and anxiety disorders: a meta-analysis of fMRI cognitive reappraisal studies. European Neuropsychopharmacology 2017 Mar;27:S78-104 [FREE Full text] [doi: 10.1016/s0924-977x(17)30150-5]
53. Chen Y, Cui Q, Xie A, Pang Y, Sheng W, Tang Q, et al. Abnormal dynamic functional connectivity density in patients with generalized anxiety disorder. J Affect Disord 2020 Jan 15;261:49-57 [FREE Full text] [doi: 10.1016/j.jad.2019.09.084] [Medline: 31600587]
54. Blair K, Aloi J, Crum K, Meffert H, White S, Taylor B, et al. Association of different types of childhood maltreatment with emotional responding and response control among youths. JAMA Netw Open 2019 May 03;2(5):e194604-e194e04 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.4604] [Medline: 31125109]
55. Goldin P, Manber-Ball T, Werner K, Heimberg R, Gross J. Neural mechanisms of cognitive reappraisal of negative self-beliefs in social anxiety disorder. Biol Psychiatry 2009 Dec 15;66(12):1091-1099 [FREE Full text] [doi: 10.1016/j.biopsych.2009.07.014] [Medline: 19717138]
56. Sakaki K, Nozawa T, Ikeda S, Kawashima R. Neural correlates of cognitive bias modification for interpretation. Soc Cogn Affect Neurosci 2020 May 11;15(2):247-260 [FREE Full text] [doi: 10.1093/scan/nsaa026] [Medline: 32322880]
57. Groenewold N, Doornbos B, Zuidersma M, Vogelzangs N, Penninx B, Aleman A, et al. Comparing cognitive and somatic symptoms of depression in myocardial infarction patients and depressed patients in primary and mental health care. PLoS One 2013;8(1):e53859 [FREE Full text] [doi: 10.1371/journal.pone.0053859] [Medline: 23342019]

Abbreviations

ANOVA: analysis of variance
BOLD: blood oxygen level dependent
fMRI: functional magnetic resonance imaging

MANOVA: multivariate analysis of variance
MRI: magnetic resonance imaging
PERS: Post-Event Rumination Scale
PRRT: Personal Relevance Rating Task
SPS: Social Phobia Scale
VR: virtual reality

Edited by J Torous; submitted 13.11.20; peer-reviewed by HK Yoon, R Dewey, M Nahum; comments to author 02.12.20; revised version received 15.01.21; accepted 12.03.21; published 14.04.21.

Please cite as:
Hur JW, Shin H, Jung D, Lee HJ, Lee S, Kim GJ, Cho CY, Choi S, Lee SM, Cho CH
Virtual Reality–Based Psychotherapy in Social Anxiety Disorder: fMRI Study Using a Self-Referential Task
JMIR Ment Health 2021;8(4):e25731
URL: https://mental.jmir.org/2021/4/e25731
doi:10.2196/25731
PMID:33851931

Original Paper

Using Machine Learning Imputed Outcomes to Assess Drug-Dependent Risk of Self-Harm in Patients with Bipolar Disorder: A Comparative Effectiveness Study

Anastasiya Nestsiarovich[1], MD, PhD; Praveen Kumar[2], MSc; Nicolas Raymond Lauve[2], BSc; Nathaniel G Hurwitz[3], MD; Aurélien J Mazurie[4], PhD; Daniel C Cannon[5], MS; Yiliang Zhu[6], PhD; Stuart James Nelson[7], MD; Annette S Crisanti[8], PhD; Berit Kerner[9], MD; Mauricio Tohen[8], MD, DrPH, MBA; Douglas J Perkins[1], PhD; Christophe Gerard Lambert[1,10], PhD

[1]Center for Global Health, Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM, United States

[2]Department of Computer Science, The University of New Mexico, Albuquerque, NM, United States

[3]New Mexico Behavioral Health Institute, Las Vegas, NM, United States

[4]TwoFoldChange Consulting, Bozeman, MT, United States

[5]Iterative Consulting, Albuquerque, NM, United States

[6]Division of Epidemiology, Biostatistics, and Preventive Medicine, Department of Internal Medicine, The University of New Mexico Health Sciences Center, Albuquerque, NM, United States

[7]Biomedical Informatics Center, George Washington University, Washington, DC, DC, United States

[8]Department of Psychiatry & Behavioral Sciences, University of New Mexico Health Sciences Center, Albuquerque, NM, United States

[9]Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[10]Division of Translational Informatics, Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM, United States

Corresponding Author:
Christophe Gerard Lambert, PhD
Center for Global Health
Department of Internal Medicine
University of New Mexico Health Sciences Center
915 Camino de Salud NE
Albuquerque, NM
United States
Phone: 1 5052729709
Email: cglambert@unm.edu

Abstract

Background: Incomplete suicidality coding in administrative claims data is a known obstacle for observational studies. With most of the negative outcomes missing from the data, it is challenging to assess the evidence on treatment strategies for the prevention of self-harm in bipolar disorder (BD), including pharmacotherapy and psychotherapy. There are conflicting data from studies on the drug-dependent risk of self-harm, and there is major uncertainty regarding the preventive effect of monotherapy and drug combinations.

Objective: The aim of this study was to compare all commonly used BD pharmacotherapies, as well as psychotherapy for the risk of self-harm, in a large population of commercially insured individuals, using self-harm imputation to overcome the known limitations of this outcome being underrecorded within US electronic health care records.

Methods: The IBM MarketScan administrative claims database was used to compare self-harm risk in patients with BD following 65 drug regimens and drug-free periods. Probable but uncoded self-harm events were imputed via machine learning, with different probability thresholds examined in a sensitivity analysis. Comparators included lithium, mood-stabilizing anticonvulsants (MSAs), second-generation antipsychotics (SGAs), first-generation antipsychotics (FGAs), and five classes of antidepressants. Cox regression models with time-varying covariates were built for individual treatment regimens and for any pharmacotherapy with or without psychosocial interventions (“psychotherapy”).

Results: Among 529,359 patients, 1.66% (n=8813 events) had imputed and/or coded self-harm following the exposure of interest. A higher self-harm risk was observed during adolescence. After multiple testing adjustment (P≤.012), the following six regimens had higher risk of self-harm than lithium: tri/tetracyclic antidepressants + SGA, FGA + MSA, FGA, serotonin-norepinephrine reuptake inhibitor (SNRI) + SGA, lithium + MSA, and lithium + SGA (hazard ratios [HRs] 1.44-2.29), and the following nine had lower risk: lamotrigine, valproate, risperidone, aripiprazole, SNRI, selective serotonin reuptake inhibitor (SSRI), "no drug," bupropion, and bupropion + SSRI (HRs 0.28-0.74). Psychotherapy alone (without medication) had a lower self-harm risk than no treatment (HR 0.56, 95% CI 0.52-0.60; $P=8.76\times10^{-58}$). The sensitivity analysis showed that the direction of drug-outcome associations did not change as a function of the self-harm probability threshold.

Conclusions: Our data support evidence on the effectiveness of antidepressants, MSAs, and psychotherapy for self-harm prevention in BD.

Trial Registration: ClinicalTrials.gov NCT02893371; https://clinicaltrials.gov/ct2/show/NCT02893371

(JMIR Ment Health 2021;8(4):e24522) doi:10.2196/24522

KEYWORDS

bipolar; mood; mania; depression; pharmacotherapy; self-harm; suicide; machine learning; psychotherapy

Introduction

Self-harming behavior is a public and mental health concern of increasing prevalence, which contributes to US hospitalization rates, morbidity, and mortality due to completed suicides. There is a clear temporal and causal link between self-injury and suicide attempts, with both being part of a "suicidality" spectrum and the former being a robust prospective predictor of the latter [1]. In 2018, suicide was the 10th leading cause of death in the general US population, reaching a rate of 14.2 per 100,000 standard population [2]. A previous study reported that the risk ratio of suicide in mental disorders was as high as 7.5 (95% CI 6.6-8.6) and in mood disorders was even higher at 12.3 (95% CI 8.9-17.1) [3]. A recent systematic review showed that bipolar disorder (BD) may be associated with the highest suicide risk among all psychiatric disorders, with over 15%-20% of deaths attributed to suicide and the standardized suicide rate being 20 to 30-fold greater than in the general population (0.2-0.4 per 100 person-years) [4]. Another review found that up to 20% of individuals with BD end their life by suicide and 20%-60% attempt suicide at least once in their lifetime [5]. The reported proportion of suicide attempts and completed suicides among individuals with BD varies from 5:1 in males over 45 years to 85:1 in females under 30 years [6].

Since suicide is an extreme form of self-harming behavior, proper recognition and management of patients presenting with self-inflicted injury are of tremendous importance to prevent lethal outcomes, especially among patients with mood disorders. The factors affecting self-harm risk should be of particular importance for studying suicidality, especially given that the self-inflicted nature of physical trauma/poisoning is often hidden owing to poor patient rapport, provider screening, and data recording.

Incomplete suicidality coding in administrative claims data is a known obstacle for observational studies. It was shown that only 19% of suicide attempts mentioned in primary care clinical notes were coded in International Classification of Diseases-9-Clinical Modification (ICD-9-CM) [7]. Our data from a large-scale observational study on imputing self-harm phenotypes in individuals with major mental illness (MMI) showed that only 1 in 19 self-harm events were coded in the billing records [8]. In addition, a methodological challenge is that ICD-9-CM coding does not robustly distinguish between suicide attempts (implying a desire to die), self-inflicted injury without suicidal intention, and suicide. While ICD-10-CM can distinguish these, many suicide attempts will be classified only under intentional self-harm. Given that all these acts are within the spectrum of self-damaging behavior, we will refer to them collectively as "self-harm." Thus, we use "self-harm"/"self-harming behavior" as the broadest term covering all forms of self-damaging acts (not thoughts alone), including not only suicide attempts, but also any intentional harm regardless of intent to die. In contrasting this self-harm study with the literature, we recognized that most of the latter was focused more narrowly on attempted and/or completed suicides.

With most of the negative outcomes missing from the data, it is challenging to assess the evidence on treatment strategies for the prevention of self-harm in BD, including pharmacotherapy and psychotherapy. There are conflicting data from studies on the drug-dependent risk of self-harm, and there is still major uncertainty regarding the preventive effect of monotherapy and drug combinations. The benefits of lowering suicidality risk were reported for lithium [9], mood-stabilizing anticonvulsants (MSAs) [10], antidepressants [11-13], and second-generation antipsychotics (SGAs) [14] in the mentally ill population. Several studies demonstrated the benefits of continuous MSA use (either alone or as an adjunct) for suicide risk reduction [15,16]. However, two recent meta-analyses showed no clear benefits of lithium [17] or valproate [18] use for preventing suicidality in patients with mood disorders. The STEP-BD study failed to find any relationship between lithium, MSA, or antipsychotic use and suicidality [19]. The US Food and Drug Administration (FDA) issued warnings for increased suicidality risk with antidepressants [20] and antiepileptic drugs [21].

Two recent meta-analyses showed that psychotherapy is associated with a reduced risk of attempting suicide, but more equivocal evidence on self-harm [22,23]; however, data on psychotherapy-dependent self-harm in adults and subjects with

BD are lacking. This provokes further questions on its relative effectiveness when compared with BD medications.

The aim of this study was to provide a comprehensive comparison of all commonly used BD pharmacotherapies, as well as psychotherapy for the risk of self-harm in a large population of commercially insured individuals, using self-harm imputation to overcome the known limitations of this outcome being underrecorded within US electronic health care record systems.

Methods

A retrospective observational study was conducted using the IBM MarketScan commercial claims and encounters (CCAE) administrative claims data and MarketScan Medicare data on 1.3 million US inpatients and outpatients with BD for the years 2003 to 2016 [24]. The database contained records of provider visits, diagnoses, procedures, outpatient prescription fills, laboratory test orders (but not results), and patient age, sex, and state of residence. The data handling was similar to that in our previous studies on the drug-dependent risk of kidney disorders and diabetes mellitus in BD [25,26], with the additional step of combining data for patients who were covered in both the CCAE and Medicare databases through their patient identifier. The relevant PostgreSQL queries and source code for data transformations and machine learning (ML) are available online [27]. The study protocol was approved by the University of New Mexico Human Research Review Committee (Institutional Review Board number 16-243).

Given that the majority of suicide attempts and self-harm events are not coded at the point of care, we employed ML to build a classification model of self-harm being present or absent, based on billing codes during emergency room (ER) or inpatient provider visits. For that purpose, we constructed a "meta-visit" by merging consecutive outpatient/inpatient/ER visits, with no gaps between visits, which allowed us to capture the medical activity associated with a given event that could have involved multiple points of care. A self-harm phenotype was defined by the presence during a meta-visit of one or more of the ICD-10-CM codes or ICD-9-CM codes listed in Multimedia Appendix 1. These encompass all codes for *intentional* self-harm or suicide attempts by any means, including poisoning. If one or more of these codes was present during a meta-visit, the meta-visit was labeled as class 1; otherwise, it was labeled as class 0.

Our earlier imputation model on over 10 million patients aged ≤65 years with MMI (schizophrenia, schizoaffective disorder, BD, and major depressive disorder) from CCAE was validated with several approaches, including via a clinician-derived "gold standard," and it identified 10.1 times more self-harm events with probability over 0.5 than were originally coded or 19 times more self-harm events based on summed probabilities [8]. In this study, we applied the previously developed ML modeling approach to an extended set of psychiatric patients of all ages, including those in the CCAE and Medicare databases. We first selected 11 million individuals with any MMI diagnosis (635,722,756 meta-visits) and performed ML on a subset of 26,392,236 meta-visits in which an inpatient or ER visit was present, using five-fold cross-validation. Covariates included age, sex, start year of the meta-visit, and the presence/absence of non–self-harm billing codes. ICD-9-CM and ICD-10-CM diagnosis codes were mapped to their Systematized Nomenclature of Medicine (SNOMED) equivalents (and all ancestors thereof) using the Observational Medical Outcomes Partnership (OMOP) vocabulary as of October 24, 2020 [28]. Procedure codes based on ICD-9-CM Volume 3 (ICD-9-CM V3), ICD-10-Procedure Coding System (ICD-10-PCS), and Current Procedural Technology, Fourth Edition (CPT-4) were mapped to ICD-10-PCS concepts (and all ancestors thereof). Overall, 190,919 covariates were added into the ML process described previously [8]. A threshold probability over 0.5 from the resulting cross-validated model estimates of self-harm was chosen to label self-harm as "present" for our main model, but sensitivity analyses were run for threshold probabilities greater than 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, and only coded self-harm (probability=1.0).

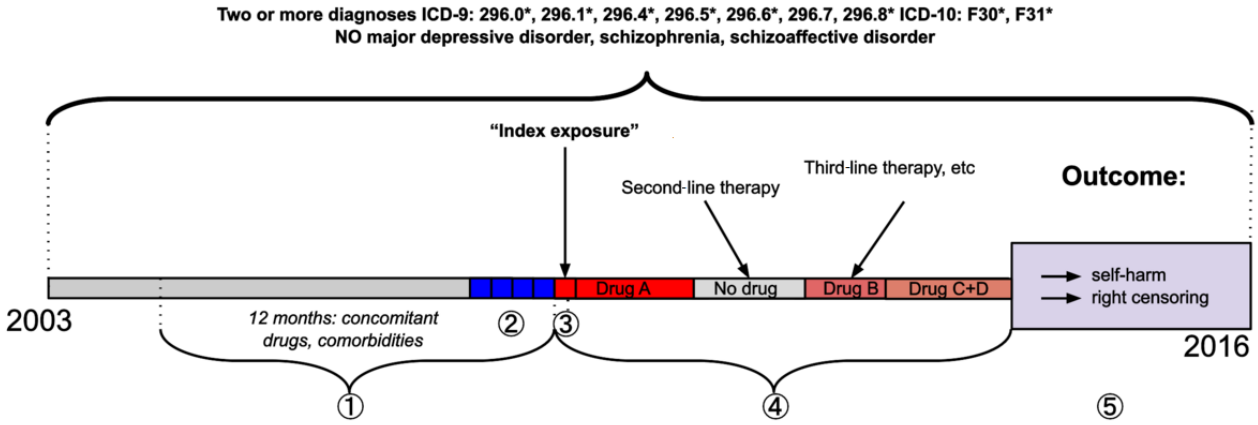
We then used the categorization of 26 million meta-visits to assign whether or not self-harm occurred following treatment exposure in a subset of 529,359 patients with two or more diagnoses of BD and no other MMI, who satisfied our data staging and inclusion/exclusion criteria (see below). Since there are approximately 28 attempts for every suicide death [29] and attempts are a subset of self-harm, selection of self-harm as the outcome allowed us to greatly increase the power of our subsequent comparative effectiveness study.

It should be noted that our ML approach was trained only on meta-visits with an inpatient/ER component since there was a negligible number of self-harm events coded during the purely outpatient meta-visits (about 1 in 100,000).

The patient inclusion criterion was two or more ICD-9-CM/ICD-10-CM diagnostic codes for BD (296.(0-1)*, 296.(4-8)*, F30*, or F31*) from 2003 to 2016. The exclusion criterion was the diagnosis of major depressive disorder, schizophrenia, or schizoaffective disorder at any time during the observation period. The onset of intellectual disability, autism spectrum disorder, mental illness of organic origin, or Parkinson disease, and use of antidementia drugs after the index exposure were considered as censoring events.

A patient was included in the analysis based on the following first observed sequence of events (Figure 1): (1) A minimum of 12 months of observation (used to compute pretreatment covariates); (2) *Index visit* (meta-visit with at least one BD diagnostic code); (3) *Index exposure* (the first day of exposure [drug regimen or "no drug"] observable on the last day of the index visit); (4) *Time-varying drug exposure period* (series of time intervals in which distinct regimens [including "no drug"] were prescribed); and (5) Outcomes of interest (the first meta-visit with newly observed coded and/or imputed self-harm and right censoring defined as any hospitalization/ER meta-visit without coded and/or imputed self-harm, or the end of patient observation).

Figure 1. Prespecified sequence of events. (1) One year before the index exposure; (2) Index visit (any meta-visit with a diagnosis of bipolar disorder); (3) Index exposure (the first day of exposure [drugs of interest or no drugs of interest] observable on the last day of the index visit); (4) Time-varying drug exposure period (series of time intervals in which distinct regimens [including "no drug"] were prescribed); (5) Outcome (the first meta-visit with coded and/or imputed self-harm or a censoring event).



The observation period ended for patients upon self-harm–unrelated hospitalization/ER meta-visit, because data on pharmacotherapy were not available during these types of visits, making it challenging to quantify psychotropic treatment time intervals. Additionally, hospitalization itself can affect the risk of self-harm.

The start and stop times were recorded for each treatment exposure period. Two Cox regression models for self-harm were built. One model compared 64 pharmacotherapies (as well as "no drug") to lithium, and the other model compared any drug (as a single category) with or without psychosocial interventions to "no treatment" (neither pharmacotherapy nor psychosocial interventions).

The idea to include "no drug" and "no treatment" in the list of comparators in our study came from patients with BD who participated in several focus groups and were engaged in designing this research [30,31]. Doing so allowed us to address patient questions regarding the safety and effectiveness of avoiding pharmacotherapy.

To ensure sufficient power to detect significant self-harm risk differences and assure convergence of Cox regression, each drug regimen was required to have 1000 or more treatment intervals and to have five or more defined cases of coded and/or imputed self-harm following exposure [32]. Because of this latter restriction, for the sensitivity analyses, the lower threshold sensitivity Cox models will have more drugs analyzed than the higher threshold ones.

The following 11 drug classes were included in the analysis: lithium, first-generation antipsychotics (FGAs), SGAs, third-generation antipsychotics (TGAs; partial agonists of dopamine receptors, aripiprazole, and brexpiprazole), MSAs, monoamine oxidase inhibitor antidepressants, noradrenergic and specific serotonergic antidepressants (NASSAs; represented by mirtazapine only), norepinephrine-dopamine reuptake inhibitors (NDRIs; represented by bupropion only), serotonin-norepinephrine reuptake inhibitors (SNRIs), selective serotonin reuptake inhibitors (SSRIs), and tri- and tetracyclic antidepressants (see Multimedia Appendix 2 for the full list of drugs).

MSAs, SGAs, and TGAs were studied as a class when used during a polypharmacy regimen exposure interval and as individual drugs when considering monotherapy time intervals. SGAs common enough for individual analysis were risperidone, olanzapine, quetiapine, ziprasidone, asenapine, paliperidone, and lurasidone. The individual MSAs studied were valproate, carbamazepine, oxcarbazepine, and lamotrigine. Of the two TGAs, only aripiprazole was common enough to be studied individually.

Combinations of two, three, or four of the 11 drug classes (represented usually by one drug from each class) with the requisite 1000 or more treatment intervals and five or more self-harm events were included in the regression model, and drug regimens without those requisites were grouped under the categories "polypharmacy 2," "polypharmacy 3," and "polypharmacy 4" (for uncommon combinations of two, three, and four or more classes, respectively). Enough instances of within-class polypharmacy were present among MSAs and SGAs to include "multi-MSA" and "multi-SGA" variables. Monotherapies without the requisite 1000 exposure intervals (clozapine, brexpiprazole, and iloperidone) were combined into the category "uncommon monotherapy."

Treatment in the main time-varying Cox regression model was represented as one or more exposure intervals, with all drug categories mutually exclusive and collectively exhaustive, using lithium monotherapy as the reference. The rules to distinguish between polypharmacy and overlapping drug regimen switch are described in our previous study of a similar design [25].

Among the covariates included in the main Cox regression model (not to be confused with the ML covariates) were patient age, sex, BD episode index visit characteristics (severity, mood polarity, and psychotic features, if documented), comorbid mental and physical conditions, including "external injury" codes evidencing noniatrogenic trauma, medication prescriptions filled (other than drugs of interest) and mental health procedures performed 1 year before (but not including) the index exposure,

hospital/ER admissions 1 year prior to the index exposure, and types of visits composing the index meta-visit (inpatient/ER/outpatient).

Patient age and the number of unique BD drugs previously tried by the patient were fitted in both Cox regression models using a smoothing spline to account for nonlinear risk of self-harm.

"Psychotherapy" included 227 procedure codes indicating psychosocial intervention (individual, group, or family psychotherapy, crisis intervention, substance abuse–focused treatment, hypnosis, biofeedback, etc) [27].

We developed two time-varying Cox regression models. In the first (main) model comparing 64 treatments and "no drug" to lithium, psychotherapy was coded as a binary time-varying covariate (indicating whether at least one of the 227 procedure codes was present during the current drug/"no drug" exposure period). In the second regression model, all drug regimens were united into a single category ("pharmacotherapy"), and psychotherapy was combined with pharmacotherapy in a time-varying covariate with the following four categories: "pharmacotherapy alone," "psychotherapy alone," "psychotherapy and pharmacotherapy," and "no psychotherapy and no pharmacotherapy" (ie, "no treatment"), with "no psychotherapy and no pharmacotherapy" as the reference.

Given the multiple treatment comparators chosen and the time-varying nature of the treatment covariates in our design, propensity score matching was not feasible for bias correction. Instead, we used a resolution IV fractional factorial design of experiments [33] (whereby main effects are aliased with three-way interactions and two-way interactions are aliased with two-way interactions) to select an appropriate subset of the 78 pretreatment covariates to control for bias. Rather than assessing whether the pretreatment covariates were associated with the outcome, we assessed in a form of sensitivity analysis whether their inclusion or exclusion impacted the hazard ratio (HR) estimates for the treatments with respect to the outcome. If so, inclusion of the variable in the model would be needed for addressing bias. If not, the variable, while possibly associated with the outcome, would nevertheless be unimportant for accurate assessment of treatment risk, and could be excluded to reduce the degrees of freedom of the model and thereby increase power. The time-varying treatment variables were included in each model, but the pretreatment covariates were included or excluded according to the factorial design across 512 different runs (plus a reference run with no pretreatment covariates) to determine which covariates had the largest impact on the drug HR coefficients. The 513 runs generated a 513×66 matrix Y of coefficients for 66 drugs over the 513 runs. The design matrix X was a 513×78 matrix of +1/−1 values corresponding to whether the given pretreatment covariate was included/excluded in a given run. Then, for each of the 66 column vectors (Y_i) of Y, a multiple linear regression was run with Y_i as the dependent variable and the 78 column vectors of X as the independent variables. We counted how many times each of the 78 covariates was significant at $P<.05/66$ over those 66 models to rank candidate covariates for our model. We discarded 26 covariates that were not significant in any of the 66 models. We then built our main Cox regression model using this set of covariates plus the treatment covariates and performed a backward elimination procedure on the pretreatment covariates, iteratively dropping the covariates that were significant in the fewest models and stopping the elimination procedure when a highly significant covariate was found (neoplasm). One drug was subsequently removed from the analysis owing to lack of events when some coding errors were corrected. We also generated an L2-norm of each row of X with the reference run row to form a vector Y' for regression with the design matrix X to assess how much the incorporation of pretreatment covariates changed all drug covariate estimates in order to understand the largest potential sources of bias. The final set of covariates selected for the first Cox model was used in the second Cox model.

The study used the following software: PostgreSQL version 10.4 (PostgreSQL Global Development Group) and R version 3.4.0 (R Foundation for Statistical Computing), including the Cox regression coxph() function from the survival (2.42-6) package and the FrF2 (1.7-2) package for fractional factorial design. All hypothesis tests were two-sided.

Results

The following self-harm classification results were observed for our MMI ML model on meta-visits based on five-fold cross-validation (probability [p] cutoff of 0.5): self-harm coded and imputed (p>0.5; N=93,311); self-harm coded but not imputed (p≤0.5; N=3717); self-harm not coded but imputed (N=1,029,058); and self-harm neither coded nor imputed (N=25,266,150) (area under the curve [AUC]=0.99; Matthews correlation coefficient [MCC]=0.28; sensitivity=0.962; specificity=0.961). The following self-harm classification results were observed when the model was applied to meta-visits for only BD cases meeting our eligibility criteria: self-harm coded and imputed (p>0.5; N=488); self-harm coded but not imputed (p≤0.5; N=37); self-harm not coded but imputed (N=8288); and self-harm neither coded nor imputed (N=520,546) (AUC=0.994; MCC=0.225; sensitivity=0.930; specificity=0.984). Thus, an extra 8288 meta-visits with imputed self-harm were added to our analytical pipeline in addition to the 525 (488+37) meta-visits that had coded self-harm for a total of 8813 persons with self-harm.

The sample sizes at different stages of the study are shown in Multimedia Appendix 3. A total of 529,359 patients met the eligibility criteria and had the prespecified sequence of events. Of them, 98.3% were censored and 1.66% (n=8813 events) had imputed and/or coded self-harm.

During the observation period after the index visit, the annual incidence of self-harm (p>0.5) was 0.013 (0.016 for all drug exposure intervals with or without psychotherapy and 0.011 for "no drug" intervals with or without psychotherapy), based on 632,512 years of observation. By summing the probabilities, during the observation period after the index visit for all exposures, the annual incidence of self-harm was 0.027 over 632,512 years of patient observation.

The 515 observed treatment regimens were collapsed to 17 monotherapies, three monoclass therapies, "no drug," and 45 drug combinations that fit the selection criteria.

The first Cox regression model comparing 65 treatment regimens to lithium showed that 11 treatments had a significantly higher risk of self-harm (P<.05, no multiple testing correction) (Table 1). The top "high-risk" treatments were "tri/tetracyclic antidepressants + SGA" (HR 2.33, 95% CI 1.28-4.26; P=5.73×10^{-3}), "SSRI + FGA" (HR 2.26, 95% CI 1.16-4.38; P=1.61×10^{-2}), "FGA + MSA" (HR 1.82, 95% CI 1.15-2.89; P=1.12×10^{-2}), and FGA monoclass therapy (HR 1.69, 95% CI 1.19-2.39; P=3.20×10^{-3}).

Nine regimens had significantly lower risk of self-harm over lithium alone (P<.05, no multiple testing correction), including monotherapies with MSAs valproate (HR 0.71, 95% CI 0.61-0.84; P=4.57×10^{-5}) and lamotrigine (HR 0.74, 95% CI 0.65-0.85; P=1.13×10^{-5}), SGAs risperidone (HR 0.68, 95% CI 0.56-0.83; P=1.82×10^{-4}) and aripiprazole (HR 0.70, 95% CI 0.59-0.84; P=9.40×10^{-5}), and antidepressant classes SNRI (HR 0.65, 95% CI 0.51-0.83; P=5.51×10^{-4}), SSRI (HR 0.61, 95% CI 0.53-0.71; P=6.05×10^{-11}), and NDRI (bupropion) (HR 0.50, 95% CI 0.39-0.65; P=1.18×10^{-7}), as well as the combination of NDRI with SSRI (HR 0.28, 95% CI 0.13-0.60; P=1.0×10^{-3}) and the "no drug" regimen.

Of the 11 polypharmacy regimens with risk significantly different from that of lithium, only bupropion + SSRI had lower risk (HR 0.28, 95% CI 0.13-0.60; P=1.00×10^{-3}). Nine of the remaining 10 high-risk polypharmacy regimens contained an antipsychotic (either SGA or FGA, or both), with the exception being lithium + MSA (HR 1.35, 95% CI 1.09-1.67; P=5.32×10^{-3}). The "no drug" exposure intervals were associated with a significantly lower risk of subsequent self-harm versus lithium monotherapy (HR 0.56, 95% CI 0.50-0.63; P=2.79×10^{-22}).

To correct for multiple comparisons, we used the Benjamini-Yekutieli procedure to reduce the false discovery rate. This correction yielded 15 regimens with a statistically significant different risk of self-harm versus lithium at a 5% false-discovery rate (which corresponded to a P value cutoff ≤.012). Six of them were of higher risk (tri/tetracyclic antidepressants + SGA, FGA + MSA, FGA, SNRI + SGA, lithium + MSA, and lithium + SGA) and nine were of lower risk than lithium (lamotrigine, valproate, risperidone, aripiprazole, SNRI, SSRI, "no drug," bupropion, and bupropion + SSRI).

Our sensitivity analysis revealed that overall most of the "high-risk" drug regimens maintained their HR values above 1 across a wide range of self-harm probability thresholds (40%-70%) (Figure 2). Only one regimen (tri/tetracyclic antidepressants + SGA) demonstrated significantly higher risk of self-harm versus lithium, across all 10 tested probability thresholds.

Table 1. Cox regression model comparing 64 pharmacotherapies and "no drug" to lithium for the risk of subsequent coded and/or imputed self-harm in patients with bipolar disorder of all ages.

Covariates[a]	HR[b,c]	Lower 95%	Upper 95%	*P* value	Patients (N=529,359)	Intervals (N=1,749,468)	Events (N=8813)
Tri/tetracyclic antidepressants + SGA[d,e]	2.33	1.28	4.26	5.73×10^{-3}	180	1044	11
SSRI[f] + FGA[g,e]	2.26	1.16	4.38	1.61×10^{-2}	195	1014	9
FGA + MSA[h,e]	1.82	1.15	2.89	1.12×10^{-2}	481	2448	19
SSRI + lithium + MSA + SGA[e]	1.72	1.00	2.97	4.96×10^{-2}	192	1424	14
FGA monoclass therapy[e]	1.69	1.19	2.39	3.20×10^{-3}	1069	5853	35
SNRI[i] + SGA[e]	1.59	1.18	2.14	2.27×10^{-3}	1466	6803	50
SNRI + MSA + SGA[e]	1.56	1.06	2.29	2.25×10^{-2}	805	4200	29
Asenapine	1.50	0.77	2.90	2.31×10^{-1}	160	1494	9
Lithium + MSA + SGA[e]	1.42	1.07	1.90	1.52×10^{-2}	1143	7748	57
Lurasidone	1.42	0.97	2.07	7.25×10^{-2}	459	3919	29
NDRI[j] + SSRI + MSA + SGA	1.40	0.66	2.96	3.84×10^{-1}	170	1251	7
SSRI + lithium + SGA	1.39	0.94	2.05	9.85×10^{-2}	725	3898	28
NDRI + lithium + MSA	1.38	0.65	2.93	3.95×10^{-1}	193	1438	7
Aripiprazole + MSA + SGA	1.38	0.86	2.22	1.87×10^{-1}	411	2641	18
Lithium + MSA[e]	1.35	1.09	1.67	5.32×10^{-3}	2763	18,728	116
Lithium + SGA[e]	1.34	1.11	1.62	2.86×10^{-3}	3757	18,980	155
Polypharmacy 4[k]	1.33	0.97	1.82	7.60×10^{-2}	1274	8867	47
SSRI + MSA + SGA[e]	1.31	1.05	1.62	1.52×10^{-2}	3153	16,322	118
NDRI + MSA + SGA	1.25	0.86	1.81	2.47×10^{-1}	836	5558	30
Aripiprazole + SGA	1.22	0.79	1.88	3.72×10^{-1}	715	3987	22
NASSA[l] + SGA	1.20	0.62	2.33	5.86×10^{-1}	321	1457	9
NDRI + lithium	1.16	0.74	1.80	5.18×10^{-1}	728	4559	21
Tri/tetracyclic antidepressants	1.15	0.71	1.88	5.66×10^{-1}	693	4130	17
SSRI + lithium + MSA	1.10	0.66	1.82	7.10×10^{-1}	499	3,514	16
Uncommon monotherapy	1.09	0.49	2.45	8.29×10^{-1}	169	1205	6
NDRI + SSRI + MSA	1.08	0.68	1.72	7.44×10^{-1}	720	4754	19
SSRI + lithium	1.06	0.81	1.39	6.61×10^{-1}	2457	13,235	64
MSA + SGA	1.05	0.91	1.21	5.15×10^{-1}	13,348	67,185	421
Lithium (reference)	1.00	N/A[m]	N/A	N/A	13,759	66,760	351
Polypharmacy 3[n]	1.00	0.79	1.25	9.66×10^{-1}	3794	23,234	95
NDRI + aripiprazole + MSA	0.99	0.51	1.91	9.67×10^{-1}	254	2316	9
Lithium + aripiprazole + MSA	0.97	0.50	1.89	9.30×10^{-1}	227	1797	9
SNRI + lithium	0.97	0.54	1.72	9.06×10^{-1}	529	2814	12
SSRI + SGA	0.96	0.80	1.15	6.36×10^{-1}	7896	33,503	188

Covariates[a]	HR[b,c]	Lower 95%	Upper 95%	*P* value	Patients (N=529,359)	Intervals (N=1,749,468)	Events (N=8813)
SSRI + MSA	0.95	0.81	1.11	5.32×10^{-1}	12,315	58,789	286
Quetiapine	0.95	0.82	1.10	5.03×10^{-1}	13,795	60,422	342
Aripiprazole + MSA	0.94	0.76	1.17	5.90×10^{-1}	3401	21,368	108
Ziprasidone	0.94	0.71	1.24	6.63×10^{-1}	2381	11,773	58
Multi-SGA	0.93	0.56	1.57	7.94×10^{-1}	677	3479	15
Lithium + aripiprazole	0.92	0.60	1.42	7.19×10^{-1}	665	4268	22
SNRI + MSA	0.90	0.68	1.20	4.80×10^{-1}	2939	13,995	58
FGA + lithium	0.89	0.40	1.99	7.73×10^{-1}	315	1472	6
NDRI + SGA	0.89	0.62	1.27	5.07×10^{-1}	1408	8336	33
Multi-MSA	0.87	0.61	1.24	4.49×10^{-1}	1451	8792	34
SSRI + aripiprazole	0.85	0.64	1.14	2.83×10^{-1}	2285	11,695	52
Olanzapine	0.84	0.66	1.07	1.57×10^{-1}	4040	16,759	83
NASSA + MSA	0.82	0.42	1.60	5.65×10^{-1}	438	2313	9
NDRI + MSA	0.82	0.64	1.05	1.20×10^{-1}	3460	21,294	74
Oxcarbazepine	0.81	0.65	1.02	6.91×10^{-2}	4436	20,633	97
Carbamazepine	0.74	0.54	1.02	6.45×10^{-2}	2284	10,662	42
Lamotrigine[e]	0.74	0.65	0.85	1.13×10^{-5}	28,624	131,786	549
NDRI + aripiprazole	0.73	0.41	1.30	2.91×10^{-1}	564	4038	12
Valproate[e]	0.71	0.61	0.84	4.57×10^{-5}	14,718	61,544	253
SSRI + aripiprazole + MSA	0.70	0.43	1.15	1.56×10^{-1}	784	5148	17
Aripiprazole[e]	0.70	0.59	0.84	9.40×10^{-5}	8872	47,373	186
Polypharmacy 2[o]	0.68	0.46	1.01	5.32×10^{-2}	2017	11,269	28
Risperidone[e]	0.68	0.56	0.83	1.82×10^{-4}	7084	28,302	138
SNRI[e]	0.65	0.51	0.83	5.51×10^{-4}	6120	27,921	78
NASSA	0.65	0.38	1.08	9.77×10^{-2}	950	4964	15
Paliperidone	0.61	0.29	1.30	2.03×10^{-1}	292	1858	7
SSRI[e]	0.61	0.53	0.71	6.05×10^{-11}	30,138	131,895	381
NDRI + SSRI + SGA	0.60	0.25	1.45	2.57×10^{-1}	345	2079	5
“No drug”[e]	0.56	0.50	0.63	2.79×10^{-22}	299,295	621,467	3694
NDRI (bupropion)[e]	0.50	0.39	0.65	1.18×10^{-7}	6005	35,433	72
SNRI + aripiprazole	0.45	0.19	1.09	7.85×10^{-2}	457	2765	5
NDRI (bupropion) +SSRI[e]	0.28	0.13	0.60	1.00×10^{-3}	1263	7496	7
Prior self-harm[e]	3.32	2.68	4.11	3.17×10^{-28}	704	1652	70
Alcohol/substance abuse or dependence[e]	1.92	1.81	2.03	7.17×10^{-106}	57,392	149,679	1065
Delirium[e]	1.69	1.36	2.10	1.92×10^{-6}	1694	4086	60

Covariates[a]	HR[b,c]	Lower 95%	Upper 95%	*P* value	Patients (N=529,359)	Intervals (N=1,749,468)	Events (N=8813)
Prior hospitalization[e]	1.63	1.54	1.72	1.78×10^{-64}	131,613	342,122	1905
Mental procedure before index exposure[e]	1.50	1.42	1.58	3.43×10^{-50}	81,943	247,474	841
Liver disease[e]	1.49	1.29	1.71	4.27×10^{-8}	9063	23,559	121
Unknown polarity of index mood episode[e]	1.33	1.24	1.43	1.67×10^{-15}	307,243	1,000,348	2447
Conduct disorder[e]	1.26	1.14	1.39	3.65×10^{-6}	13,728	39,657	255
Seizure disorder[e]	1.18	1.11	1.25	2.07×10^{-7}	90,276	307,887	667
External injury[e]	1.12	1.06	1.19	2.31×10^{-5}	117,722	338,635	1350
Pulmonary disorder	1.12	0.99	1.26	7.60×10^{-2}	18,960	48,974	160
Depression during the index meta-visit[e]	1.10	1.02	1.18	1.52×10^{-2}	72,386	245,707	487
Male sex[e]	1.07	1.03	1.12	1.46×10^{-3}	227,507	733,963	1966
Exposure to sedative or antianxiety drug[e]	1.06	1.00	1.12	4.24×10^{-2}	130,186	431,040	929
Number of prior unique BD[p] drugs tried (linear component of spline fit)[e]	1.02	1.01	1.04	4.17×10^{-3}	N/A	N/A	N/A
Psychotic features present during the index meta-visit	1.01	0.92	1.11	7.73×10^{-1}	23,846	78,579	345
Age (linear component of spline fit)[e]	0.98	0.97	0.98	6.99×10^{-201}	N/A	N/A	N/A
Exposure to central nervous system stimulant	0.95	0.88	1.01	9.62×10^{-2}	55,140	195,116	396
BD type II during the index meta-visit[e]	0.93	0.87	0.99	2.49×10^{-2}	90,741	331,720	497
Manic episode during the index meta-visit[e]	0.92	0.85	0.99	3.40×10^{-2}	82,955	250,285	512
Exposure to glucocorticoids[e]	0.83	0.77	0.89	1.04×10^{-7}	80,626	246,182	472
Exposure to antibacterial agents[e]	0.80	0.76	0.84	3.77×10^{-17}	144,933	453,129	951
Exposure to sex hormones[e]	0.79	0.74	0.85	3.08×10^{-10}	67,550	217,335	372
Neoplasm[e]	0.77	0.70	0.85	1.70×10^{-7}	42,628	130,488	220
Psychotic features unknown during the index meta-visit[e]	0.67	0.63	0.72	2.85×10^{-30}	401,789	1,295,408	2709
Psychotherapy (psychosocial interventions)[e]	0.59	0.57	0.62	1.12×10^{-114}	249,328	704,937	1369
Outpatient visit present during the index meta-visit[e]	0.56	0.50	0.63	4.69×10^{-23}	522,232	1,732,715	3475
Age (nonlinear components of the spline model)[e]	N/A	N/A	N/A	4.58×10^{-32}	N/A	N/A	N/A

Covariates[a]	HR[b,c]	Lower 95%	Upper 95%	*P* value	Patients (N=529,359)	Intervals (N=1,749,468)	Events (N=8813)
Number of prior unique BD drugs tried (non-linear components of spline fit)[e]	N/A	N/A	N/A	1.21×10^{-9}	N/A	N/A	N/A

[a]Covariates labeled "prior" are related to the 1-year period before the index exposure.

[b]Covariates are sorted by their hazard ratio value.

[c]HR: hazard ratio.

[d]SGA: second-generation antipsychotic.

[e]Covariates with significant *P* values (<.05; no multiple testing correction).

[f]SSRI: selective serotonin reuptake inhibitor.

[g]FGA: first-generation antipsychotic.

[h]MSA: mood stabilizing anticonvulsant.

[i]SNRI: serotonin-norepinephrine reuptake inhibitor.

[j]NDRI: norepinephrine-dopamine reuptake inhibitor (represented by bupropion only).

[k]Polypharmacy 4: uncommon combination of four or more bipolar disorder drug classes.

[l]NASSA: noradrenergic and specific serotonergic antidepressant (represented by mirtazapine only).

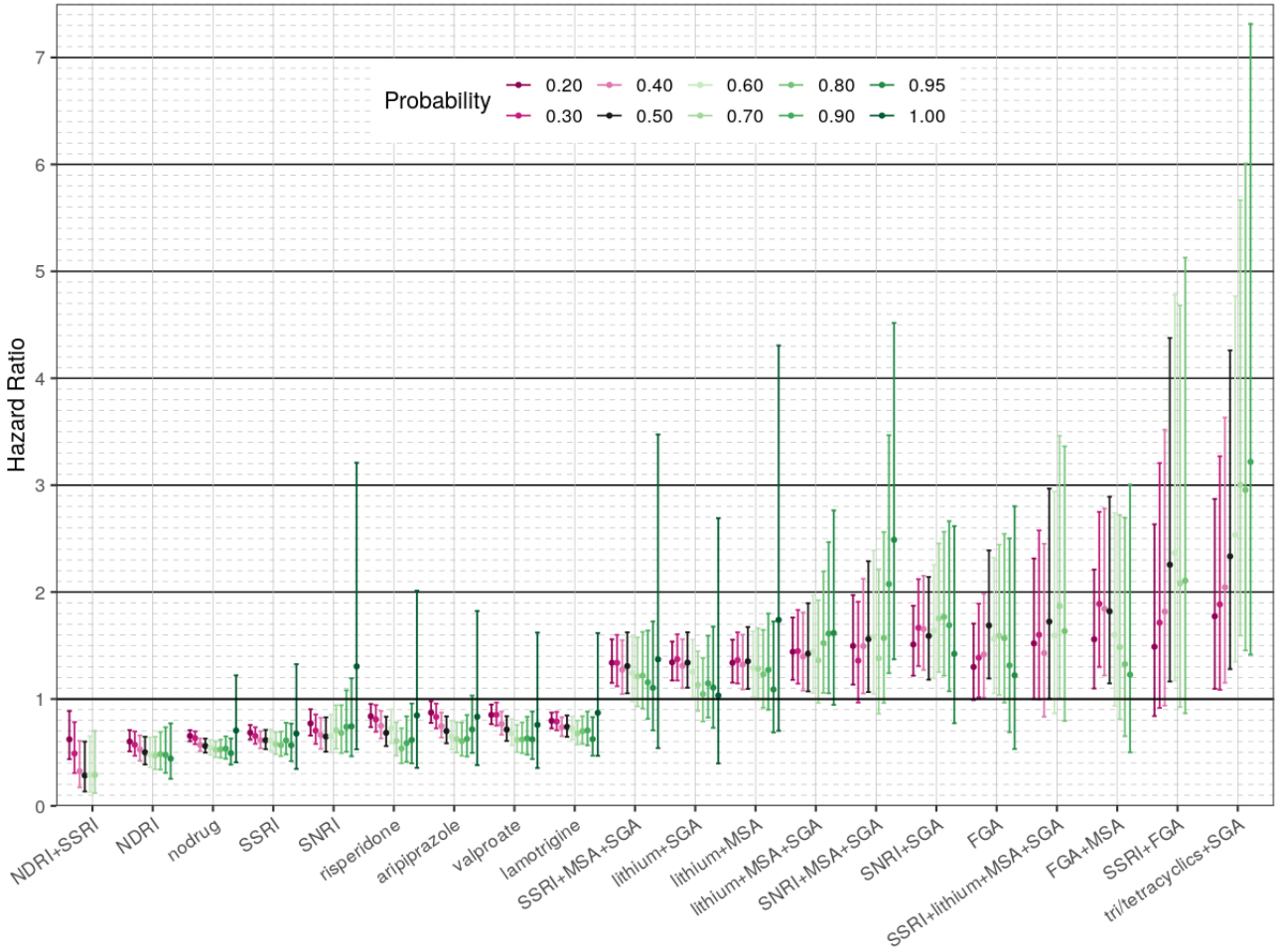
[m]N/A: not applicable.

[n]Polypharmacy 3: uncommon combination of three bipolar disorder drug classes.

[o]Polypharmacy 2: uncommon combination of two bipolar disorder drug classes.

[p]BD: bipolar disorder.

Figure 2. Sensitivity analysis for the "low-risk" and "high-risk" covariates in the first regression model comparing individual exposure regimens for the risk of self-harm. The X-axis shows 13 covariates and the respective 20%-100% self-harm thresholds chosen to impute the outcome. The Y-axis shows the respective hazard ratios (colored dots) and 95% CIs (colored lines). Varied intensity magenta is used to represent the range of 20%-40% self-harm probability thresholds, black is used to represent the 50% threshold of the main model, and varied intensity green is used to represent the 60%-100% probability threshold used. Missing estimates are due to lack of sufficient outcomes for a regimen to be included (observed in the higher probability threshold models). MSA: mood-stabilizing anticonvulsant; NDRI: norepinephrine-dopamine reuptake inhibitor; nodrug: period free from any of the studied bipolar disorder drugs; SGA: second-generation antipsychotic; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor.



For most of the "low-risk" drugs, the HR values were below 1 at any self-harm probability threshold, except for 90%-100% (very likely to be self-harm or actually coded). Bupropion alone or in combination with SSRI had a significant association with lower self-harm risk across all tested thresholds. As expected, the higher the probability of self-harm, the larger were the respective HR CIs owing to fewer events observed. The results of the sensitivity analysis for all exposure covariates in this model, as well as the nondrug covariates, can be found in Multimedia Appendix 4, Multimedia Appendix 5, Multimedia Appendix 6, and Multimedia Appendix 7.

When assessing the largest sources of bias, four variables were highly significantly associated with shifting the estimates of all the treatment coefficients, based on the regression of *Y'* versus *X*, including the number of prior unique BD drugs tried, psychotherapy, alcohol/substance abuse or dependence, and outpatient visit present during the index meta-visit. These four were among the covariates with the top five most significant (P<.05/66) associations over the 66 Y_i versus X regressions performed on individual treatment estimates in our variable selection procedure. A total of 29 pretreatment covariates were incorporated in the model to adjust for potential bias in treatment risk estimates.

In the main Cox model, documentation of prior coded self-harm had the highest HR value among all nondrug covariates (HR 3.32, 95% CI 2.68-4.11; $P=3.17\times10^{-28}$). A set of mental conditions, including delirium, substance/alcohol abuse and dependence, conduct disorder, and procedures related to mental health services were associated with a significantly higher risk of self-harm (HR 1.26-1.92, P<.05). Previous hospitalizations, liver disease, and seizures were also associated with elevated self-harm risk when present (HR 1.18-1.63, P<.05). Exposure to antianxiety and sedative drugs showed a modest risk of self-harm (HR 1.06, 95% CI 1.00-1.12; $P=4.24\times10^{-2}$). Additionally, index visit depression was modestly associated with self-harm risk (HR 1.10, 95% CI 1.02-1.18; $P=1.52\times10^{-2}$).

Multiple factors had significantly lower self-harm risk, including index manic mood episodes, BD type II, use of antibacterial agents and glucocorticoids, exposure to sex hormones, and neoplasm diagnosis (HR 0.77-0.93; P<.05). Psychotherapy during the exposure period was strongly associated with a lower

risk of self-harm (HR 0.59, 95% CI 0.57-0.62; P=1.12×10^{-114}). The lowest HR value for self-harm was associated with an outpatient visit being present during the index meta-visit (HR 0.56, 95% CI 0.50-0.63; P=4.69×10^{-23}) (Table 1).

When self-harm risk was plotted as a function of the number of different unique drugs of interest tried in the past, we observed that HR values slightly decreased after intervals with one and two drugs used, but then started to rise with the number of agents used (Figure 3).

Figure 4 shows the risk of self-harm as a function of patient age. It demonstrates that HR values were much higher in adolescence, dropped after the 20s, and leveled off with older age.

Figure 3. The hazard ratio of coded and/or imputed self-harm as a function of the number of different unique drugs of interest used by the patient in the year prior to the index visit plus up to the prior treatment interval. The graph represents a smoothing spline, with the reference being zero prior drugs. The blue dotted lines represent 95% CIs.

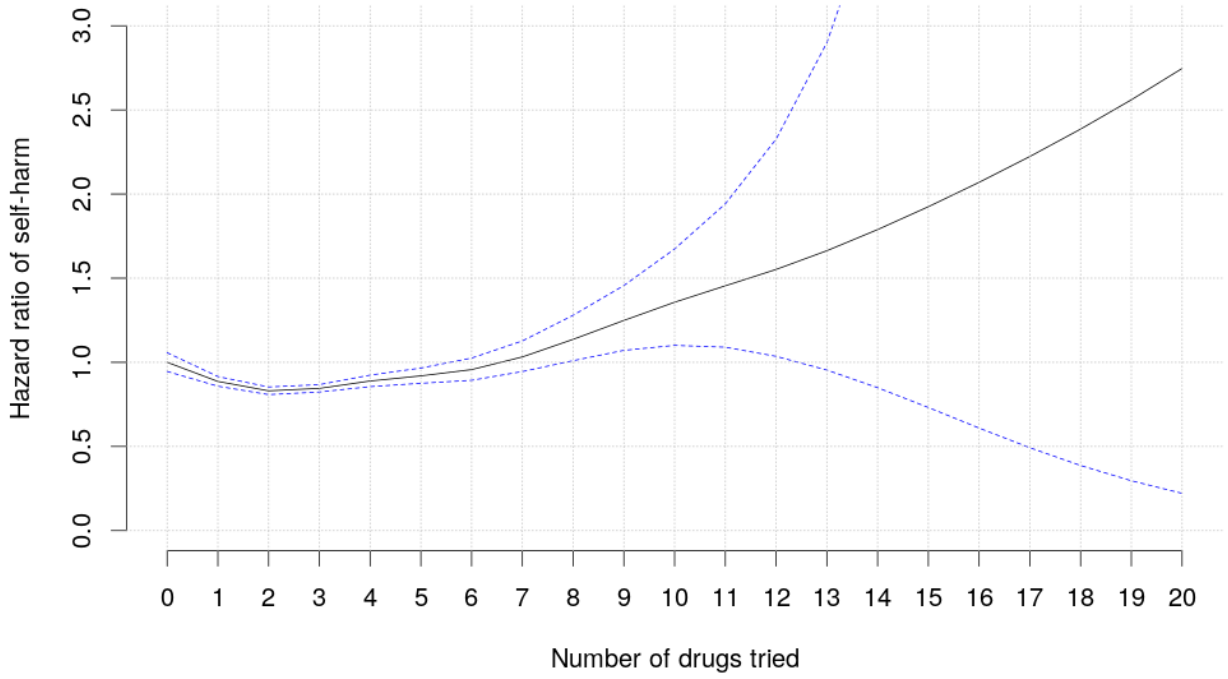
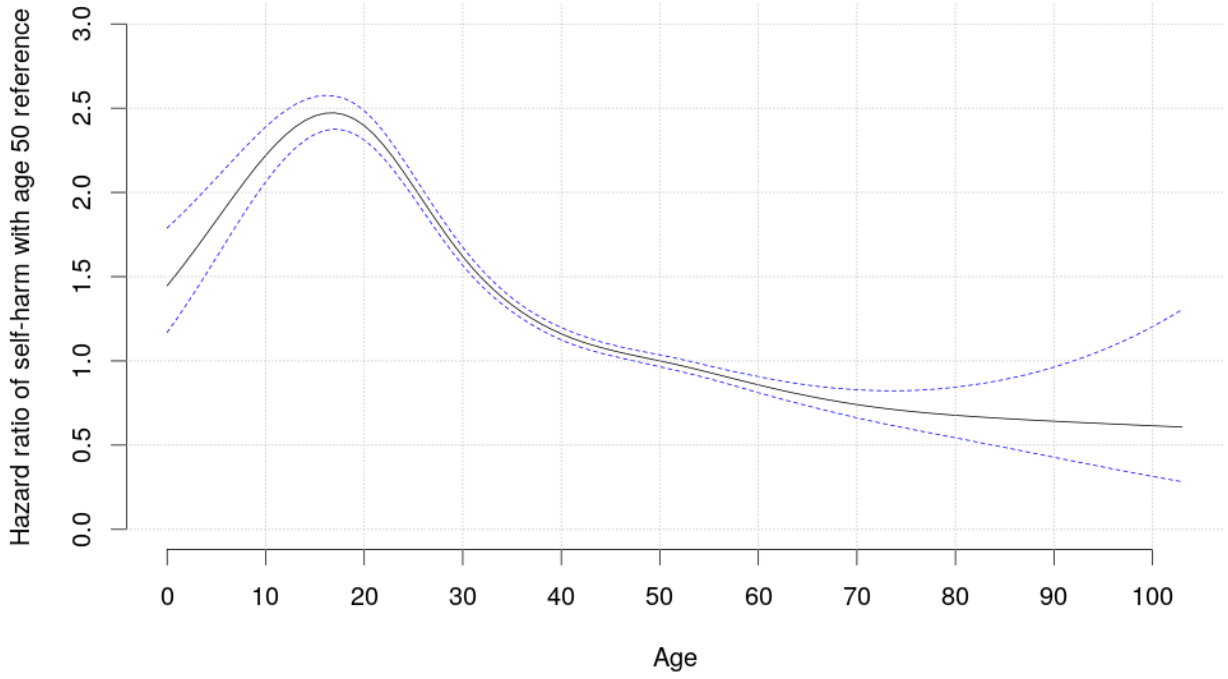


Figure 4. The hazard ratio of coded and/or imputed self-harm as a function of patient age. The graph represents a smoothing spline, with the reference being age 50 years. The blue dotted lines represent 95% CIs.



In the second Cox regression model with all BD drugs grouped under the "pharmacotherapy" category, the risk of self-harm was the lowest following "psychotherapy alone" intervals, compared with "no treatment" (HR 0.56, 95% CI 0.52-0.60; P=8.76×10^{-58}) (Table 2). The combination "psychotherapy + pharmacotherapy" had a somewhat lower risk of self-harm (HR 0.88, 95% CI 0.83-0.95; P=3.80×10^{-4}), but pharmacotherapy alone was associated with a significantly higher risk compared with "no treatment" (HR 1.38, 95% CI 1.30-1.48; P=1.09×10^{-22}).

Table 2. Cox regression model comparing "pharmacotherapy" (as a single category) and "psychotherapy" (psychosocial interventions) to "no treatment" (no drugs and no psychotherapy) for the risk of subsequent coded and/or imputed self-harm in patients with bipolar disorder of all ages.

Covariates[a]	HR[b,c]	Lower 95%	Upper 95%	*P* value	Patients (N=529,359)	Intervals (N=1,749,468)	Events (N=8813)
Pharmacotherapy alone (any drug regimen)[d]	1.38	1.30	1.48	1.09×10^{-22}	122,898	651,191	2819
Any drug and psychotherapy[d]	0.88	0.83	0.95	3.80×10^{-4}	107,166	476,810	2300
Psychotherapy alone[d]	0.56	0.52	0.60	8.76×10^{-58}	142,162	228,127	1183
No treatment (reference)	1.00	N/A[e]	N/A	N/A	157,133	393,340	2511
Prior self-harm[d]	3.32	2.68	4.10	3.34×10^{-28}	704	1652	70
Alcohol/substance abuse or dependence[d]	1.91	1.80	2.03	2.48×10^{-105}	57,392	149,679	1065
Delirium[d]	1.68	1.35	2.08	2.74×10^{-6}	1694	4086	60
Previous hospitalization[d]	1.61	1.52	1.71	3.21×10^{-62}	131,613	342,122	1905
Prior mental health procedure[d]	1.49	1.41	1.57	2.90×10^{-49}	81,943	247,474	841
Liver disease[d]	1.48	1.29	1.71	4.30×10^{-8}	9063	23,559	121
Unknown polarity of index mood episode[d]	1.35	1.25	1.45	2.33×10^{-16}	307,243	1,000,348	2447
Conduct disorder[d]	1.25	1.13	1.38	9.03×10^{-6}	13,728	39,657	255
Seizure disorder[d]	1.17	1.10	1.25	4.56×10^{-7}	90,276	307,887	667
Pulmonary disorder	1.12	0.99	1.26	6.69×10^{-2}	18,960	48,974	160
External injury	1.12	1.06	1.18	4.79×10^{-5}	117,722	338,635	1350
Depression during the index visit[d]	1.09	1.01	1.17	3.15×10^{-2}	72,386	245,707	487
Male sex[d]	1.08	1.04	1.13	3.96×10^{-4}	227,507	733,963	1966
Number of prior unique BD[f] drugs tried (linear component of spline fit)[d]	1.07	1.06	1.09	2.99×10^{-24}	N/A	N/A	N/A
Exposure to sedative antianxiety	1.05	0.99	1.11	1.10×10^{-1}	130,186	431,040	929
Psychotic features present	1.03	0.93	1.13	6.01×10^{-1}	23,846	78,579	345
Age (linear component of spline fit)[d]	0.98	0.98	0.98	5.17×10^{-194}	N/A	N/A	N/A
Exposure to central nervous system stimulant[d]	0.93	0.87	0.99	2.15×10^{-2}	55,140	195,116	396
BD type II during the index meta-visit[d]	0.92	0.86	0.98	6.72×10^{-3}	90,741	331,720	497
Manic episode during the index meta-visit[d]	0.91	0.84	0.98	1.45×10^{-2}	82,955	250,285	512
Exposure to glucocorticoids[d]	0.82	0.77	0.88	2.18×10^{-8}	80,626	246,182	472
Exposure to antibacterial agents[d]	0.78	0.74	0.83	2.65×10^{-19}	144,933	453,129	951
Exposure to sex hormones[d]	0.78	0.73	0.84	3.77×10^{-11}	67,550	217,335	372
Neoplasm[d]	0.76	0.69	0.84	4.76×10^{-8}	42,628	130,488	220
Psychotic features unknown during the index meta-visit[d]	0.66	0.62	0.71	8.75×10^{-33}	401,789	1,295,408	2709
Outpatient visit present during the index meta-visit[d]	0.56	0.50	0.62	3.19×10^{-24}	522,232	1,732,715	3475
Age (nonlinear components of spline model)[d]	N/A	N/A	N/A	3.00×10^{-33}	N/A	N/A	N/A
Number of prior unique BD drugs tried (nonlinear components of spline fit)[d]	N/A	N/A	N/A	4.47×10^{-12}	N/A	N/A	N/A

[a]Covariates labeled "prior" are related to the 1-year period before the index exposure.

[b]Covariates are sorted by their hazard ratio value.

[c]HR: hazard ratio.

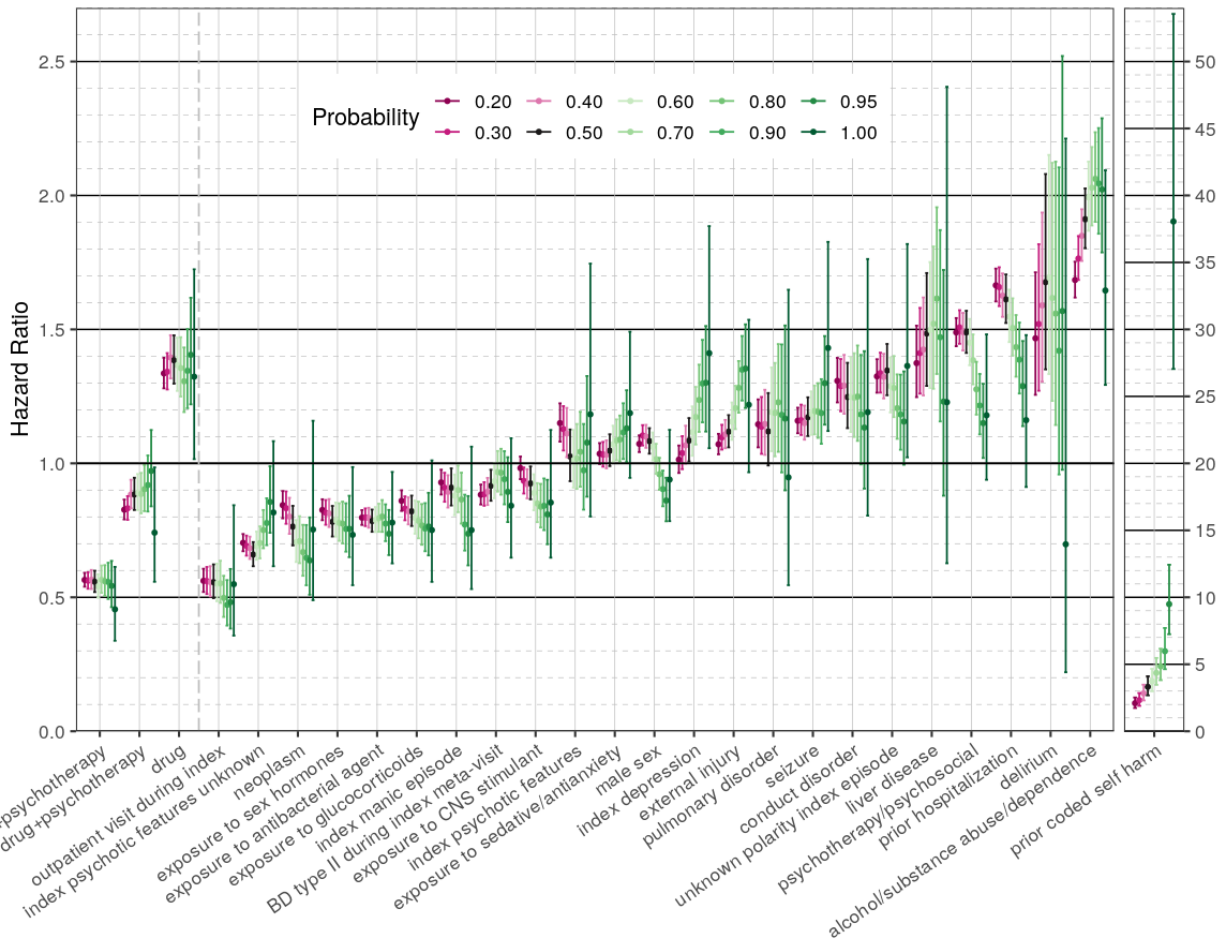
[d]Covariates with significant *P* values (<.05).

[e]N/A: not applicable.

[f]BD: bipolar disorder.

The sensitivity analysis showed that most of the "high-risk" variables maintained their HR values above 1 at a wide range of self-harm probability thresholds, except for very high thresholds (>80%-90%). Prior self-harm and pharmacotherapy alone (without psychosocial interventions) had significantly high HR values across all tested self-harm probability thresholds. The "low-risk" variables mostly maintained their HR values below 1 with different self-harm probability thresholds (except for 80%-100%). Five variables had significantly lower risk of self-harm across all tested thresholds compared with no treatment at all. They were psychotherapy alone, prior self-harm, outpatient visit present during the index meta-visit, exposure to sex hormones, and use of antibacterial agents. As in the first model, the higher was the probability of self-harm, the wider were the CIs owing to fewer events (Figure 5).

Figure 5. Sensitivity analysis for the "low-risk" and "high-risk" covariates in the second regression model comparing pharmacotherapy (as a single exposure category) and psychotherapy for the risk of self-harm. The X-axis shows 27 covariates and the respective 20%-100% self-harm probability thresholds chosen to impute the outcome. The Y-axis shows the respective hazard ratios (HRs) (colored dots) and CIs (colored lines). Varied intensity magenta is used to represent the range of 20%-40% self-harm probability thresholds, black is used to represent the 50% threshold of the main model, and varied intensity green is used to represent the 60%-100% probability threshold used. The covariate "prior coded self-harm" is separated out with a different HR scale in the far right, since the HR values were extremely high at the 100% (coded) probability threshold. BD: bipolar disorder; CNS: central nervous system; Drug: any of the bipolar disorder drugs of interest.



Discussion

Principal Findings

Given the use of imputed self-harm (in addition to formally coded) as the primary outcome in our study, it is worthwhile to compare the coded and imputed annual incidence of self-harm in our data with that of the literature. In a recent UK study [34], within 25,965 person-years of observation in a cohort of 6671 patients with pharmacologically treated BD, who were aged 16 years or above, the annual incidence of hospitalized self-harm was 3774 per 100,000 person-years at risk (PYAR). The coded self-harm in our BD cohort of all ages was only 83 per 100,000 PYAR. This would constitute 1:45-fold underrecording, if US

rates of self-harm are comparable to UK rates. Our earlier estimate [8] that only 1 in 19 self-harm events was coded (within meta-visits having an inpatient and/or ER component) may not have been sufficiently pessimistic. In contrast to the strikingly low rates of coded self-harm in our study data, the estimates of coded + imputed self-harm used for our main model were more reassuring, with 1393 self-harm events per 100,000 PYAR. When summing the probabilities over all meta-visits, our estimate of the level of self-harm was 2839 per 100,000 PYAR, which is 75% of the UK estimate and is probably still low. Our sensitivity analysis revealed a range of 525 (formally coded only) to 20,226 (coded + imputed with >20% probability) self-harm events corresponding to a range of 83 to 3198/100,000 PYAR. It is important to note that because HRs are relative measures, they may be stably estimated across a broad range of imputation thresholds, with the advantage of more power for lower thresholds.

Our findings suggest that exposure to FGAs and some multidrug combinations were associated with 1.31 to 2.33 higher risks of self-harm compared with lithium; however, these associations were possibly observed owing to multiple-testing type I error. Drug-free intervals ("no drug") had one of the lowest HR values in our first regression model compared with lithium (HR 0.56, 95% CI 0.50-0.63; P=2.79×10^{-22}). According to a recent literature review, there is strong converging evidence indicating that long-term lithium treatment lowers deaths by suicide in patients with BD [4], which can be attributed to its possible serotonergic effect [35]. One explanation for the better performance of the "no drug" regimen in our study versus lithium could be indication bias, as drug-free periods can be associated with stable remission or asymptomatic states.

Self-harm risk reduction was significant with monotherapies involving the MSAs valproate and lamotrigine, the atypical antipsychotics risperidone and aripiprazole, the antidepressant bupropion, and monoclass treatment with SNRI and SSRI antidepressants. There are conflicting data in the literature on antidepressant-dependent suicidality in mood disorders, with reports on both the increased [36] and decreased risks of suicidal behavior [11,37]. In 2004, antidepressants received an FDA black box warning owing to increased suicidal thoughts and behaviors in adolescents on antidepressants versus placebo in FDA approval–seeking trials [20], and this warning was extended to include young adults in 2007 [38]. It is still not entirely clear whether a presumed increased suicidality risk in antidepressant users is due to drugs failing to prevent deterioration involving the natural illness course, due to their activating effect, or due to manic switch with subsequent mood phase inversion. In contrast, a 27-year prospective study on mood disorders showed that the risk of suicide attempts or suicides was reduced by 20% among participants taking antidepressants (HR 0.80, 95% CI 0.68-0.95; P=.011) [12]. Subsequent findings of the same authors showed that suicidality risk was reduced by 54% in individuals with BD type I and by 35% in those with BD type II while on antidepressants, compared with propensity-matched unexposed intervals [13]. While our study generated evidence on a more broadly defined set of "self-harm" acts, our data support the findings of the relative safety of SSRI and SNRI antidepressants compared with lithium and even with "no drug" in relation to suicidality in BD.

SGAs were previously shown to be associated with a reduced risk of suicide in patients with schizophrenia [14], although two recent international observational studies demonstrated the inferiority of quetiapine and olanzapine compared with lithium for self-harm prevention [34], and even an increased risk of completed suicide among BD patients taking antipsychotics [39]. Our data showed that the SGA risperidone is associated with a significantly lower self-harm risk in patients with BD. There is evidence suggesting that the beneficial effect of antipsychotics in BD may be explained by reduced impulsivity and risk taking [40].

Similar to studies on antidepressants, there are conflicting data on the MSA-dependent risk of self-harm in BD. While some studies reported an equally beneficial effect of MSAs (divalproex and carbamazepine) to lithium for BD suicidality prevention [10], others reported a significantly safer profile for lithium [41]. The majority of MSAs received an FDA warning of increased suicidal thoughts and behaviors in 2008 [42], based on a meta-analysis of 11 drugs [21]. Several studies failed to find any significant changes in suicidality risk according to antiepileptic drug intake [18,37]. However, a large pharmacoepidemiologic study found significantly lower rates of suicide attempts following MSA use, compared with the period before treatment, and showed that MSA monotherapy was significantly protective relative to no pharmacologic treatment (3 per 1000 vs 15 per 1000 person-years) [43]. Our findings support the evidence of a beneficial role of MSAs in self-harm prevention in BD management. Unlike the other data [34], our data showed that valproate is superior to lithium in terms of the association with reduced self-harm risk.

Given that 10 of the 11 "high-risk" exposures in our study were polypharmacy regimens, we made efforts to address the possible indication bias of multidrug regimens being given to patients who are treatment-resistant, by modeling the risk of self-harm as a function of the number of unique BD drugs filled in the year prior to the index visit plus those drugs tried from the index visit up to the current treatment interval. We fit this within the Cox regression model using a smoothing spline with no prior drugs set as the reference (Figure 3). The risk of self-harm was significantly lower in individuals treated with one to five different BD drugs in the year prior to the index visit, compared with individuals who had no prior drugs in the observed period of time. One explanation for this finding is that several "trial and error" attempts eventually result in better control over illness symptoms. However, self-harm risk was significantly higher in patients who received eight or more unique BD drugs, compared with drug-naive subjects, evidencing drug-resistant cases. At the same time, the rapidly expanding range of 95% CI corresponding to 8 to 20 drugs indicates limited sample sizes in this range. Overall, given that our self-harm risk estimates for the drugs account for prior treatment complexity and that the magnitude of this factor's impact on risk was modest, it seems unlikely that a presumed polypharmacy-dependent increase in self-harm risk in patients with BD is fully explained by drug resistance or disease severity. However, we may not have fully corrected for indication biases. In particular, we did

not model drug exposures prior to the year before the index visit.

Our sensitivity analysis showed that the direction of drug-outcome associations did not change as a function of the threshold of self-harm probability, while HR CIs were much more narrow when the outcome was imputed rather than coded. This provides evidence that using ML-imputed outcomes is a promising approach to increase power to perform comparative effectiveness studies, particularly when a phenotype is sparsely coded.

The presence of an outpatient encounter during the index meta-visit (with or without an adjacent hospitalization/ER visit) was associated with the lowest risk of self-harm. This can be explained by more accessible or comprehensive health care services provided, as evidenced by a patient visiting his/her outpatient provider during a crisis.

As was expected, our data suggested that psychosocial interventions may decrease the risk of self-harm in patients with BD. A recent meta-analysis showed that patients who received psychotherapy were less likely to subsequently attempt suicide [22]. However, a surprising finding from our second Cox regression model was that the HR of self-harm was lower following time intervals with psychotherapy alone, rather than when psychotherapy was combined with pharmacotherapy. This could be explained by indication bias, since drug-free patients could be asymptomatic or in stable remission. Another explanation is that pharmacotherapy was a very heterogeneous category combining “low-risk” and “high-risk” regimens together. There was insufficient power to perform a per-drug analysis of adjunctive psychotherapy.

The study limitations include nonrandomized assignment of patients to treatment groups; no patient data availability prior to the insurance enrollment date, as well as prior to 2003; unmeasured indication or other biases (eg, personality traits, coping strategies, environmental stressors, and support systems); and no correction for medication dosage, route of administration, or release mechanism.

Conclusions

The risk of self-harm varied more than eight-fold among different BD drug regimens. Exposure to antidepressant or MSA monotherapy was associated with a significantly lower risk of subsequent self-harm compared with lithium. Psychotherapy was strongly associated with a decreased risk of self-harm in patients with BD. ML imputation of self-harm can enhance the power for comparative effectiveness studies of BD treatments. The risk of self-harm was the highest during adolescence. Our data support the evidence that prior self-harm is one of the strongest predictors of future self-harm.

Acknowledgments

This work was supported by the Patient-Centered Outcomes Research Institute (PCORI) (award CER-1507-3160) and was part of the research project “Longitudinal comparative effectiveness of bipolar disorder therapies” (ClinicalTrials.gov identifier: NCT02893371). The study protocol was approved by the University of New Mexico Human Research Review Committee (Institutional Review Board number 16-243). PCORI funding was used for database access and salaries for all members of the research team who, in turn, contributed to study design, data collection, analysis and interpretation, writing of the article, and the decision to submit the article for publication. The views in this publication are solely the responsibility of the authors and do not necessarily represent the views of PCORI, its Board of Governors, or the Methodology Committee. We dedicate this article to our dear friend and coauthor, Dr Nathaniel G Hurwitz, who passed away during its preparation. He worked on this project as a labor of love, both for us and for all those affected by bipolar disorder.

Authors' Contributions

AN: study design, data analysis and interpretation, and manuscript writing; PK: study design, data analysis, and manuscript editing; NRL: study design and data analysis; NGH: study design and data interpretation; AJM: data extraction and analysis; DCC: data extraction and analysis; YZ: data analysis and interpretation, and manuscript editing; SJN: study design, data interpretation, and manuscript editing; ASC: study design and manuscript editing; BK: study design, data analysis and interpretation, and manuscript editing; MT: study design, data interpretation, and manuscript editing; DJP: study design, data interpretation, and manuscript editing; CGL: study coordination, study design, data analysis and interpretation, and manuscript writing and editing.

Conflicts of Interest

MT was a full-time employee at Lilly (1997 to 2008). He has received honoraria from or consulted for Abbott, Actavis, AstraZeneca, Bristol Myers Squibb, GlaxoSmithKline, Lilly, Johnson & Johnson, Otsuka, Merck, Sunovion, Forest, Gedeon Richter, Roche, Elan, Alkermes, Allergan, Lundbeck, Teva, Pamlab, Wyeth, and Wiley Publishing. His spouse was a full-time employee at Lilly (1998-2013). AN, AJM, NGH, BK, YZ, SJN, ASC, DJP, CGL, DCC, PK, and NRL report no relevant financial relationships with commercial interests.

Multimedia Appendix 1
List of self-harm billing codes.
[DOCX File , 13 KB - mental_v8i4e24522_app1.docx]

Multimedia Appendix 2
Drugs of interest analyzed either individually or as part of a composite class.
[DOCX File , 16 KB - mental_v8i4e24522_app2.docx]

Multimedia Appendix 3
Sample size based on data staging. BD: bipolar disorder; ER: emergency room; MMI: major mental illness.
[PNG File , 100 KB - mental_v8i4e24522_app3.png]

Multimedia Appendix 4
Sensitivity analysis for the first third of drug covariates (sorted by their hazard ratio from low to high) in the regression model comparing individual exposure regimens for the risk of self-harm. The X-axis shows drug covariates and the respective 20%-100% self-harm thresholds chosen to impute the outcome. The Y-axis shows the respective hazard ratios (colored dots) and CIs (colored lines). Varied intensity magenta is used to represent the range of 20%-40% self-harm probability thresholds, black is used to represent the 50% threshold of the main model, and varied intensity green is used to represent the 60%-100% probability threshold used. MSA: mood-stabilizing anticonvulsant; NDRI: norepinephrine-dopamine reuptake inhibitor; SGA: second-generation antipsychotic; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor.
[PNG File , 121 KB - mental_v8i4e24522_app4.png]

Multimedia Appendix 5
Sensitivity analysis for the second third of drug covariates (sorted by their hazard ratio from low to high) in the regression model comparing individual exposure regimens for the risk of self-harm. The X-axis shows drug covariates and the respective 20%-100% self-harm thresholds chosen to impute the outcome. The Y-axis shows the respective hazard ratios (colored dots) and CIs (colored lines). Varied intensity magenta is used to represent the range of 20%-40% self-harm probability thresholds, black is used to represent the 50% threshold of the main model, and varied intensity green is used to represent the 60%-100% probability threshold used. MSA: mood-stabilizing anticonvulsant; NDRI: norepinephrine-dopamine reuptake inhibitor; SGA: second-generation antipsychotic; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor.
[PNG File , 133 KB - mental_v8i4e24522_app5.png]

Multimedia Appendix 6
Sensitivity analysis for the last third of drug covariates (sorted by their hazard ratio from low to high) in the regression model comparing individual exposure regimens for the risk of self-harm. The X-axis shows drug covariates and the respective 20%-100% self-harm thresholds chosen to impute the outcome. The Y-axis shows the respective hazard ratios (colored dots) and CIs (colored lines). Varied intensity magenta is used to represent the range of 20%-40% self-harm probability thresholds, black is used to represent the 50% threshold of the main model, and varied intensity green is used to represent the 60%-100% probability threshold used. MSA: mood-stabilizing anticonvulsant; NDRI: norepinephrine-dopamine reuptake inhibitor; SGA: second-generation antipsychotic; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor.
[PNG File , 160 KB - mental_v8i4e24522_app6.png]

Multimedia Appendix 7
Sensitivity analysis for the nondrug covariates in the regression model comparing individual exposure regimens for the risk of self-harm. The X-axis shows nondrug covariates and the respective 20%-100% self-harm thresholds chosen to impute the outcome. The Y-axis shows the respective hazard ratios (HRs) (colored dots) and CIs (colored lines). Varied intensity magenta is used to represent the range of 20%-40% self-harm probability thresholds, black is used to represent the 50% threshold of the main model, and varied intensity green is used to represent the 60%-100% probability threshold used. The covariate “prior coded self-harm” is separated out with a different HR scale in the far right, since the HR values were extremely high at 100% (coded) probability threshold. MSA: mood-stabilizing anticonvulsant; NDRI: norepinephrine-dopamine reuptake inhibitor; SGA: second-generation antipsychotic; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor.
[PNG File , 146 KB - mental_v8i4e24522_app7.png]

References

1. Cooper J, Kapur N, Webb R, Lawlor M, Guthrie E, Mackway-Jones K, et al. Suicide after deliberate self-harm: a 4-year cohort study. Am J Psychiatry 2005 Feb;162(2):297-303. [doi: 10.1176/appi.ajp.162.2.297] [Medline: 15677594]
2. Xu J, Murphy SL, Kochanek KD, Arias E. Mortality in the United States, 2018. Centers for Disease Control and Prevention. 2020 Jan. URL: https://www.cdc.gov/nchs/data/databriefs/db355-h.pdf [accessed 2020-02-08]
3. Too LS, Spittal MJ, Bugeja L, Reifels L, Butterworth P, Pirkis J. The association between mental disorders and suicide: A systematic review and meta-analysis of record linkage studies. J Affect Disord 2019 Dec 01;259:302-313 [FREE Full text] [doi: 10.1016/j.jad.2019.08.054] [Medline: 31450139]

4. Plans L, Barrot C, Nieto E, Rios J, Schulze TG, Papiol S, et al. Association between completed suicide and bipolar disorder: A systematic review of the literature. J Affect Disord 2019 Jan 01;242:111-122. [doi: 10.1016/j.jad.2018.08.054] [Medline: 30173059]
5. Dome P, Rihmer Z, Gonda X. Suicide Risk in Bipolar Disorder: A Brief Review. Medicina (Kaunas) 2019 Jul 24;55(8) [FREE Full text] [doi: 10.3390/medicina55080403] [Medline: 31344941]
6. Simon GE, Hunkeler E, Fireman B, Lee JY, Savarino J. Risk of suicide attempt and suicide death in patients treated for bipolar disorder. Bipolar Disord 2007 Aug;9(5):526-530. [doi: 10.1111/j.1399-5618.2007.00408.x] [Medline: 17680924]
7. Anderson HD, Pace WD, Brandt E, Nielsen RD, Allen RR, Libby AM, et al. Monitoring suicidal patients in primary care using electronic health records. J Am Board Fam Med 2015;28(1):65-71 [FREE Full text] [doi: 10.3122/jabfm.2015.01.140181] [Medline: 25567824]
8. Kumar P, Nestsiarovich A, Nelson SJ, Kerner B, Perkins DJ, Lambert CG. Imputation and characterization of uncoded self-harm in major mental illness using machine learning. J Am Med Inform Assoc 2020 Jan 01;27(1):136-146 [FREE Full text] [doi: 10.1093/jamia/ocz173] [Medline: 31651956]
9. Baldessarini RJ, Tondo L, Davis P, Pompili M, Goodwin FK, Hennen J. Decreased risk of suicides and attempts during long-term lithium treatment: a meta-analytic review. Bipolar Disord 2006 Oct;8(5 Pt 2):625-639. [doi: 10.1111/j.1399-5618.2006.00344.x] [Medline: 17042835]
10. Yerevanian BI, Koek RJ, Mintz J. Bipolar pharmacotherapy and suicidal behavior. Part I: Lithium, divalproex and carbamazepine. J Affect Disord 2007 Nov;103(1-3):5-11. [doi: 10.1016/j.jad.2007.05.019] [Medline: 17628692]
11. Angst F, Stassen HH, Clayton PJ, Angst J. Mortality of patients with mood disorders: follow-up over 34–38 years. Journal of Affective Disorders 2002 Apr;68(2-3):167-181. [doi: 10.1016/s0165-0327(01)00377-9] [Medline: 12063145]
12. Leon AC, Solomon DA, Li C, Fiedorowicz JG, Coryell WH, Endicott J, et al. Antidepressants and risks of suicide and suicide attempts: a 27-year observational study. J Clin Psychiatry 2011 May;72(5):580-586 [FREE Full text] [doi: 10.4088/JCP.10m06552] [Medline: 21658345]
13. Leon AC, Fiedorowicz JG, Solomon DA, Li C, Coryell WH, Endicott J, et al. Risk of suicidal behavior with antidepressants in bipolar and unipolar disorders. J Clin Psychiatry 2014 Jul;75(7):720-727 [FREE Full text] [doi: 10.4088/JCP.13m08744] [Medline: 25093469]
14. Reutfors J, Bahmanyar S, Jönsson EG, Brandt L, Bodén R, Ekbom A, et al. Medication and suicide risk in schizophrenia: a nested case-control study. Schizophr Res 2013 Nov;150(2-3):416-420. [doi: 10.1016/j.schres.2013.09.001] [Medline: 24094723]
15. Søndergård L, Lopez AG, Andersen PK, Kessing LV. Mood-stabilizing pharmacological treatment in bipolar disorders and risk of suicide. Bipolar Disord 2008 Feb;10(1):87-94. [doi: 10.1111/j.1399-5618.2008.00464.x] [Medline: 18199245]
16. Smith EG, Søndergård L, Lopez AG, Andersen PK, Kessing LV. Association between consistent purchase of anticonvulsants or lithium and suicide risk: a longitudinal cohort study from Denmark, 1995-2001. J Affect Disord 2009 Oct;117(3):162-167. [doi: 10.1016/j.jad.2009.01.013] [Medline: 19243837]
17. Cipriani A, Hawton K, Stockton S, Geddes JR. Lithium in the prevention of suicide in mood disorders: updated systematic review and meta-analysis. BMJ 2013 Jun 27;346:f3646 [FREE Full text] [doi: 10.1136/bmj.f3646] [Medline: 23814104]
18. Chen T, Kamali M, Chu C, Yeh C, Huang S, Mao W, et al. Divalproex and its effect on suicide risk in bipolar disorder: A systematic review and meta-analysis of multinational observational studies. J Affect Disord 2019 Feb 15;245:812-818. [doi: 10.1016/j.jad.2018.11.093] [Medline: 30699864]
19. Marangell LB, Dennehy EB, Wisniewski SR, Bauer MS, Miyahara S, Allen MH, et al. Case-control analyses of the impact of pharmacotherapy on prospectively observed suicide attempts and completed suicides in bipolar disorder: findings from STEP-BD. J Clin Psychiatry 2008 Jun;69(6):916-922. [doi: 10.4088/jcp.v69n0605] [Medline: 18399724]
20. Suicidality in Children and Adolescents Being Treated With Antidepressant Medications. US Food and Drug Administration. 2018. URL: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/suicidality-children-and-adolescents-being-treated-antidepressant-medications [accessed 2020-06-05]
21. Statistical review and evaluation: Antiepileptic drugs and suicidality. US Food and Drug Administration. 2008. URL: https://www.fda.gov/files/drugs/published/Statistical-Review-and-Evaluation--Antiepileptic-Drugs-and-Suicidality.pdf [accessed 2020-06-08]
22. Calati R, Courtet P. Is psychotherapy effective for reducing suicide attempt and non-suicidal self-injury rates? Meta-analysis and meta-regression of literature data. J Psychiatr Res 2016 Aug;79:8-20. [doi: 10.1016/j.jpsychires.2016.04.003] [Medline: 27128172]
23. Briggs S, Netuveli G, Gould N, Gkaravella A, Gluckman NS, Kangogyere P, et al. The effectiveness of psychoanalytic/psychodynamic psychotherapy for reducing suicide attempts and self-harm: systematic review and meta-analysis. Br J Psychiatry 2019 Jun;214(6):320-328. [doi: 10.1192/bjp.2019.33] [Medline: 30816079]
24. Quint JB. Health research data for the real world: the MarketScan databases. Ann Arbor, MI: Truven Health Analytics; Jan 2015.
25. Nestsiarovich A, Kerner B, Mazurie AJ, Cannon DC, Hurwitz NG, Zhu Y, et al. Comparison of 71 bipolar disorder pharmacotherapies for kidney disorder risk: The potential hazards of polypharmacy. J Affect Disord 2019 Jun 01;252:201-211 [FREE Full text] [doi: 10.1016/j.jad.2019.04.009] [Medline: 30986735]

26. Nestsiarovich A, Kerner B, Mazurie AJ, Cannon DC, Hurwitz NG, Zhu Y, et al. Diabetes mellitus risk for 102 drugs and drug combinations used in patients with bipolar disorder. Psychoneuroendocrinology 2020 Feb;112:104511 [FREE Full text] [doi: 10.1016/j.psyneuen.2019.104511] [Medline: 31744781]
27. Self Harm Pharmacotherapy Psychotherapy. GitLab. URL: https://gitlab.com/PCORIUNMPUBLIC/self_harm_pharmacotherapy_psychotherapy [accessed 2021-04-09]
28. Athena – OHDSI vocabularies repository. OHDSI – Observational Health Data Sciences and Informatics. URL: https://athena.ohdsi.org/search-terms/start [accessed 2021-04-08]
29. Suicide statistics. American Foundation for Suicide Prevention. 2019. URL: https://afsp.org/suicide-statistics/ [accessed 2020-06-08]
30. Nestsiarovich A, Hurwitz NG, Nelson SJ, Crisanti AS, Kerner B, Kuntz MJ, et al. Systemic challenges in bipolar disorder management: A patient-centered approach. Bipolar Disord 2017 Dec;19(8):676-688 [FREE Full text] [doi: 10.1111/bdi.12547] [Medline: 28901625]
31. Kerner B, Crisanti AS, DeShaw JL, Ho JG, Jordan K, Krall RL, et al. Preferences of Information Dissemination on Treatment for Bipolar Disorder: Patient-Centered Focus Group Study. JMIR Ment Health 2019 Jun 25;6(6):e12848 [FREE Full text] [doi: 10.2196/12848] [Medline: 31237566]
32. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007 Mar 15;165(6):710-718. [doi: 10.1093/aje/kwk052] [Medline: 17182981]
33. Margolin BH. Resolution Iv Fractional Factorial Designs. Journal of the Royal Statistical Society: Series B (Methodological) 2018 Dec 05;31(3):514-523. [doi: 10.1111/j.2517-6161.1969.tb00811.x]
34. Hayes JF, Pitman A, Marston L, Walters K, Geddes JR, King M, et al. Self-harm, Unintentional Injury, and Suicide in Bipolar Disorder During Maintenance Mood Stabilizer Treatment: A UK Population-Based Electronic Health Records Study. JAMA Psychiatry 2016 Jun 01;73(6):630-637. [doi: 10.1001/jamapsychiatry.2016.0432] [Medline: 27167638]
35. Kovacsics CE, Gottesman II, Gould TD. Lithium's antisuicidal efficacy: elucidation of neurobiological targets using endophenotype strategies. Annu Rev Pharmacol Toxicol 2009;49:175-198. [doi: 10.1146/annurev.pharmtox.011008.145557] [Medline: 18834309]
36. Yerevanian BI, Koek RJ, Mintz J, Akiskal HS. Bipolar pharmacotherapy and suicidal behavior Part 2. The impact of antidepressants. J Affect Disord 2007 Nov;103(1-3):13-21. [doi: 10.1016/j.jad.2007.05.017] [Medline: 17617467]
37. Leon AC, Demirtas H, Li C, Hedeker D. Two propensity score-based strategies for a three-decade observational study: investigating psychotropic medications and suicide risk. Stat Med 2012 Nov 30;31(27):3255-3260. [doi: 10.1002/sim.5339] [Medline: 22865578]
38. Stone M, Laughren T, Jones ML, Levenson M, Holland PC, Hughes A, et al. Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. BMJ 2009 Aug 11;339:b2880 [FREE Full text] [doi: 10.1136/bmj.b2880] [Medline: 19671933]
39. Takeuchi T, Takenoshita S, Taka F, Nakao M, Nomura K. The Relationship between Psychotropic Drug Use and Suicidal Behavior in Japan: Japanese Adverse Drug Event Report. Pharmacopsychiatry 2017 Mar;50(2):69-73. [doi: 10.1055/s-0042-113468] [Medline: 27595297]
40. Reddy LF, Lee J, Davis MC, Altshuler L, Glahn DC, Miklowitz DJ, et al. Impulsivity and risk taking in bipolar disorder and schizophrenia. Neuropsychopharmacology 2014 Jan;39(2):456-463 [FREE Full text] [doi: 10.1038/npp.2013.218] [Medline: 23963117]
41. Goodwin FK, Fireman B, Simon GE, Hunkeler EM, Lee J, Revicki D. Suicide risk in bipolar disorder during treatment with lithium and divalproex. JAMA 2003 Sep 17;290(11):1467-1473. [doi: 10.1001/jama.290.11.1467] [Medline: 13129986]
42. Mittal M, Harrison DL, Miller MJ, Farmer KC, Thompson DM, Ng Y. Have antiepileptic drug prescription claims changed following the FDA suicidality warning? An evaluation in a state Medicaid program. Epilepsy Behav 2014 May;34:109-115 [FREE Full text] [doi: 10.1016/j.yebeh.2014.03.017] [Medline: 24739448]
43. Gibbons RD, Hur K, Brown CH, Mann JJ. Relationship between antiepileptic drugs and suicide attempts in patients with bipolar disorder. Arch Gen Psychiatry 2009 Dec;66(12):1354-1360 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2009.159] [Medline: 19996040]

Abbreviations

AUC: area under curve
BD: bipolar disorder
CCAE: commercial claims and encounters
ER: emergency room
FDA: US Food and Drug Administration
FGA: first-generation antipsychotic
HR: hazard ratio
ICD: International Classification of Diseases
MCC: Matthews correlation coefficient

ML: machine learning
MMI: major mental illness
MSA: mood-stabilizing anticonvulsant
NASSA: noradrenergic and specific serotonergic antidepressant
NDRI: norepinephrine-dopamine reuptake inhibitor
PYAR: person-years at risk
SGA: second-generation antipsychotic
SNRI: serotonin-norepinephrine reuptake inhibitor
SSRI: selective serotonin reuptake inhibitor
TGA: third-generation antipsychotic

Edited by J Torous; submitted 23.09.20; peer-reviewed by G Simon, R Jacobucci; comments to author 14.12.20; revised version received 08.02.21; accepted 09.03.21; published 21.04.21.

Please cite as:
Nestsiarovich A, Kumar P, Lauve NR, Hurwitz NG, Mazurie AJ, Cannon DC, Zhu Y, Nelson SJ, Crisanti AS, Kerner B, Tohen M, Perkins DJ, Lambert CG
Using Machine Learning Imputed Outcomes to Assess Drug-Dependent Risk of Self-Harm in Patients with Bipolar Disorder: A Comparative Effectiveness Study
JMIR Ment Health 2021;8(4):e24522
URL: https://mental.jmir.org/2021/4/e24522
doi:10.2196/24522
PMID:33688834

Original Paper

Effectiveness of Delivering Dialectical Behavioral Therapy Techniques by Email in Patients With Borderline Personality Disorder: Nonrandomized Controlled Trial

Nazanin Alavi[1,2], MD, FRCPC; Callum Stephenson[1,2], BScH; Margo Rivera[3], PhD

[1]Department of Psychiatry, Queen's University, Kingston, ON, Canada

[2]Centre for Neuroscience Studies, Faculty of Health Sciences, Queen's University, Kingston, ON, Canada

[3]Personality Disorder Services, Queen's University, Kingston, ON, Canada

Corresponding Author:
Nazanin Alavi, MD, FRCPC
Department of Psychiatry
Queen's University
Hotel Dieu Hospital
166 Brock Street
Kingston, ON, K7L 5G2
Canada
Phone: 1 613 544 3310
Email: nazanin.alavitabari@kingstonhsc.ca

Abstract

Background: Borderline personality disorder is a debilitating and prevalent mental health disorder, with often inaccessible treatment options. Electronically delivered dialectical behavioral therapy could be an efficacious and more accessible intervention.

Objective: We aimed to evaluate the efficacy of electronic delivery of dialectical behavioral therapy in the treatment of individuals with symptoms of borderline personality disorder.

Methods: Study participants diagnosed with borderline personality disorder were offered either an email-based or in-person group format dialectical behavioral therapy skill-building program. During each session, participants were provided with both the material and feedback regarding their previous week's homework. Electronically delivered dialectical behavioral therapy protocol and content were designed to mirror in-person content. Participants were assessed using the Self-Assessment Questionnaire (SAQ) and Difficulties in Emotion Regulation Scale (DERS).

Results: There were significant increases in SAQ scores from pre- to posttreatment in the electronic delivery group ($F_{1,92}$=69.32, P<.001) and in-person group ($F_{1,92}$=60.97, P<.001). There were no significant differences observed between the groups at pre- and posttreatment for SAQ scores ($F_{1,92}$=.05, P=.83). There were significant decreases in DERS scores observed between pre- and posttreatment in the electronic delivery group ($F_{1,91}$=30.15, P<.001) and the in-person group ($F_{1,91}$=58.18, P<.001). There were no significant differences observed between the groups for DERS scores pre- and posttreatment ($F_{1,91}$=.24, P=.63). There was no significant difference in treatment efficacy observed between the 2 treatment arms (P<.001).

Conclusions: Despite the proven efficacy of in-person dialectical behavioral therapy in the treatment of borderline personality disorder, there are barriers to receiving this treatment. With the prevalence of internet access continuing to rise globally, delivering dialectical behavioral therapy with email may provide a more accessible alternative to treatment for individuals with borderline personality disorder without sacrificing the quality of care.

Trial Registration: ClinicalTrials.gov NCT04493580; https://clinicaltrials.gov/ct2/show/NCT04493580

(JMIR Ment Health 2021;8(4):e27308) doi:10.2196/27308

KEYWORDS

borderline personality disorder; treatment; psychotherapy; psychotherapy; dialectical behavioral therapy; barriers to treatment; mental health; online; internet; electronic

Introduction

The essential features of a personality disorder are an impairment in personality (self and interpersonal) functioning and the presence of pathological personality traits, including impairments in identity, self-direction, and interpersonal functioning [1]. Additionally, it is common for an individual with a personality disorder to present with numerous comorbid mental disorder diagnoses [2]. Moreover, it is well documented that the presence of a personality disorder negatively impacts the efficacy of treatment for other physical and mental disorders [1,3,4].

Borderline personality disorder is a serious psychiatric disorder with a prevalence of approximately 1% to 2% in the general population [5]. Borderline personality disorder is characterized by a pervasive pattern of mental instability in the areas of affect regulation, impulse control, interpersonal relationships, and self-image. Clinical signs of borderline personality disorder include emotional dysregulation, impulsive aggression, repeated intentional self-injury, and chronic suicidal tendencies and ideation [6].

Dialectical behavior therapy (DBT) is one of the most efficacious and commonly utilized modalities for the treatment of borderline personality disorder [7]. DBT is structured with weekly group meetings where participants can learn about and develop coping skills. This learning and development are done using an evidence-based manualized curriculum that encompasses the topics of mindfulness, interpersonal relationships, emotion regulation, and distress tolerance [7,8].

Although various psychotherapy modalities have been proven to be effective in treating borderline personality disorder, they are often not used due to limited resources and lack of accessibility for patients. The causes of inaccessibility to treatment can be broken into 2 categories—practical, such as long waitlists, rural or remote living situations where treatment is not available, and lack of transportation, and psychological, wherein individuals with borderline personality disorder are more resistant to the idea of participating in group therapy than individuals with other mental disorders [9]. A commonly reported reason for avoiding in-person psychotherapy is to avoid the stigma surrounding mental health [10]. Due to these factors, while potentially effective, in-person DBT does not appear to be an ideal treatment modality for individuals with borderline personality disorder, and the issues mentioned must be addressed in the hopes of developing a more accessible treatment option.

A viable treatment delivery method may be using the internet. Internet usage is increasing exponentially, with over 2.5 billion people globally using the internet [11]. Moreover, there was a reported increase of 56% in internet users globally from 2000 to 2012 [11]. Currently, approximately 41% of households worldwide can connect to the internet, and approximately 37% of women and 41% of men use the internet [11]. For individuals in high-income economies, internet use has become an integral part of daily life. Even within middle- and low-income countries, internet usage continues to increase [11]. With higher speeds, more affordable access, and an increasing user base, there is a growing demand for more robust and sophisticated technologies and applications on the internet [11].

Given the culture shift with respect to internet communication, it is not surprising that there has been rapid growth in recent years in the research, development, and use of internet-based psychotherapeutic interventions. To address the issue of the abovementioned barriers to access, the use of internet-based psychotherapy, which is clinically effective, has emerged as a solution [12]. In-person and live participation in psychotherapy is no longer the exclusive treatment delivery route for individuals to address their mental health needs. Fortunately, research has shown that electronically delivered cognitive behavioral therapy (e-CBT) is a cost-effective and easily accessible treatment modality for a wide variety of mental health disorders [13,14]. Furthermore, it has been suggested that internet-based psychotherapy can increase treatment adherence and yield high treatment satisfaction among patients while offering results comparable to those offered by in-person psychotherapy. For the treatment of depression, e-CBT programs have demonstrated considerable efficacy and have been increasingly utilized to enhance access to care for individuals [15,16]. Although there has been a large amount of research on the efficacy of e-CBT in the treatment of mental health disorders, to date, no study has examined the efficacy of electronically delivered DBT (e-DBT) skill-building programs for treating individuals with borderline personality disorder.

We aimed to add to the literature by creating and offering an email-based DBT skill-building program as an alternative treatment modality for individuals with borderline personality disorder who were referred to participate in in-person DBT. Additionally, we aimed to explore the efficacy and accessibility of e-DBT compared to in-person DBT.

Methods

Dialectical Behavioral Therapy

Since 1995, the Personality Disorders Service in Kingston, Ontario, Canada, has developed psychotherapeutic programs that have integrated a range of modalities for the treatment of individuals with borderline personality disorder [17,18]. Currently, several therapy groups are offered, including a weekly skill-building group structured to offer the basic DBT curriculum, titled *Managing Powerful Emotions*. This curriculum includes mindfulness, distress tolerance, emotion regulation, and interpersonal effectiveness and has been offered since 2000 as a first-line treatment for individuals diagnosed with borderline personality disorder.

The Personality Disorders Service offers more advanced therapy groups for individuals who have completed the Managing Powerful Emotions program and wish to continue seeking treatment modalities. One of the most intensive therapy programs offered is the Chrysalis Day Treatment Program [17,18]. To participate in the Chrysalis Day Treatment Program, an individual must progress through 2 prior phases: (1) Managing Powerful Emotions and (2) psychotherapy groups incorporating DBT skill-building. The Chrysalis Day Treatment Program is an intensive day treatment program that integrates

DBT skill-building, psychodynamic psychotherapy, and a range of other group therapy modalities. This integrated form of psychotherapy is extremely effective, particularly in individuals with more severe and prolonged symptomology and trauma histories [18].

Recruitment

Individuals who were referred to the Personality Disorders Service in Kingston, Ontario, Canada (after confirmation of diagnosis of borderline personality disorder by a psychiatrist in the Department of Psychiatry at Queen's University) were offered the opportunity to select either the in-person DBT skill-building program or an email format of the program. Inclusion criteria were being between the ages of 18 and 65 years at study inception and a diagnosis of borderline personality disorder according to the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition guidelines. Moreover, participants were required to have the competency to consent and participate, the ability to speak and read English, and to have consistent and reliable access to the internet. Participants were excluded from the study if they were experiencing acute hypomanic or manic episodes, were experiencing acute psychosis, had severe alcohol or substance use disorders, or were currently receiving DBT.

Individuals who were referred to the program were provided with an information sheet with details of the study and the comparative effectiveness of online and in-person treatment. Individuals were asked to give informed consent (ie, sign a letter of consent) to participate in the study. The in-person treatment group served as a control group.

Measurement Scales

All participants were required to complete questionnaires at baseline, at the end of week 7, and after the completion of the treatment program. These questionnaires included the Self-Assessment Questionnaire (SAQ) and the Difficulties in Emotion Regulation Scale (DERS) [19]. The DERS is a self-report tool designed to obtain an overall measure of the difficulty respondents have with various aspects of emotion regulation. The DERS provides an overall score of difficulties with emotion regulation as well as an assessment of each of the following 6 specific factors related to emotion dysregulation: nonacceptance (nonacceptance of emotional responses), goals (difficulty engaging in goal-oriented behaviors), impulse (difficulty controlling impulses), awareness (lack of emotional awareness), strategies (lack of access to emotion regulation strategies), and clarity (lack of emotional clarity).

Participants were informed that both in-person and e-DBT treatment programs were created with the intent of helping them to learn useful skills and strategies for managing emotions and behaviors and that it was not to be used as a crisis service. Participants of the electronically delivered program were informed that their therapist would read their emails once a week and would not be able to respond to crises, such as acute suicidal ideation or intent. Participants were informed that, in the case of an emergency, they should either go to their local emergency department or call emergency services or their local crisis line.

Therapy Programs

The study protocol was registered (NCT04493580). Both programs had a duration of 15 weeks, with 1 DBT session per week. In the e-DBT group, participants were individually emailed approximately 30 to 40 PowerPoint slides (Microsoft Inc) each week that they were to complete. These slides included general information on a particular topic (Table 1), an overview of skills related to the topics being covered, and homework sheets to be completed and returned to their therapists. The team of therapists involved in care delivery were psychiatry residents, psychologists, and registered nurses who also facilitated the in-person groups. All content and the format of the e-DBT program were designed to directly correspond with those of the in-person group.

Participants in the e-DBT program were asked to email their homework sheets back to their therapist by a specific day each week. The following day, the therapist would review the homework and email the participant individualized feedback regarding their homework along with the following week's homework and slides. To be eligible to receive the following week's materials, participants were required to email their homework before the set deadline. If the homework was not submitted before the deadline, a reminder email was sent. If a participant missed more than 2 sessions, they were removed from the study.

Table 1. Managing Powerful Emotions sessions.

Week	Content
1	Goals, Accepting Reality, Willingness vs. Willfulness, Distress Tolerance Box
2	Crisis Survival Strategies – Distract
3	Crisis Survival Strategies – Self-Soothe
4	Crisis Survival Strategies – Improve the Moment
5	Crisis Survival Strategies – Pros and Cons
6	Skills for Accepting Life as it is in the Moment – Observing Your Breath
7	Skills for Accepting Life as it is in the Moment – Half-Smiling Exercises
8	Skills for Accepting Life as it is in the Moment – Awareness Exercises
9	Distress Tolerance Box Due / Emotion Regulation
10	Myths About Emotion
11	Model for Describing Emotion, Emotion Sheet
12	Functions of Emotion
13	Reducing Vulnerability to Painful Emotions, Increasing Pleasant Events
14	Acting Opposite to Action Urge
15	Your Opinions

Analysis

To determine whether there was a significant change in functioning or level of symptomatology pre- to posttreatment, *t* tests and mixed-model analysis of variance 2 (e-DBT, in-person) ×2 (pretreatment, posttreatment) with Bonferroni correction were used. Between group and within-group differences at baseline were assessed using *t* tests.

Ethical Considerations and Confidentiality

Only individuals involved in the direct care of participants had access to their information. Data regarding study variables were entered anonymously into a database separate from clinical files using anonymous participant identification numbers. This study was approved by the Queen's University Research and Ethics Board (TRAQ 6007697; PSIY-391-13).

Results

Participants

Of 107 individuals recruited for the study, 52 elected to take part in the e-DBT group (male n=10, female n=42), and 55 elected to be in the in-person group (male n=14, female n=41) (Figure 1). At baseline, there were no significant differences between the 2 groups in SAQ (e-DBT: mean 27.29, SD 6.36; in-person: mean 27.45, SD 48.00; t_{105}=–.11, *P*=.91) or DERS scores (e-DBT: mean 52.88, SD 23.45; in-person: mean 57.81, SD 21.51; t_{104}=–1.13, *P*=.26) (Figure 2). Of the participants in the e-DBT group, 23 completed all therapy sessions, with those who did not complete all sessions completing between 2 and 13 sessions (mean 8.83, SD 3.30). Of the participants in the in-person group, 27 completed all therapy sessions, with those not completing all sessions completing between 1 and 14 sessions (mean 4.29, SD 3.46).

Figure 1. Participant enrollment, allocation, and analysis process. e-DBT: electronically delivered dialectical behavioral therapy.

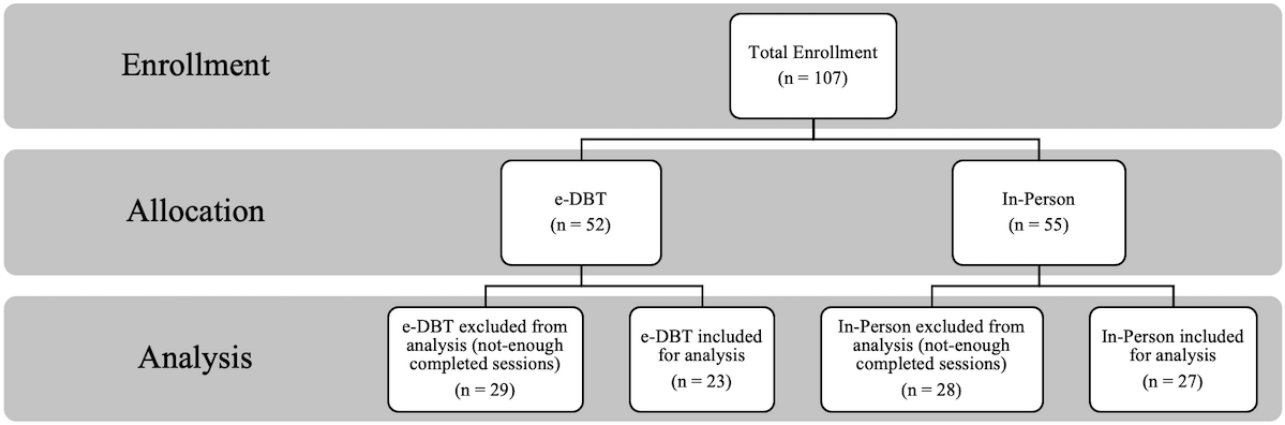
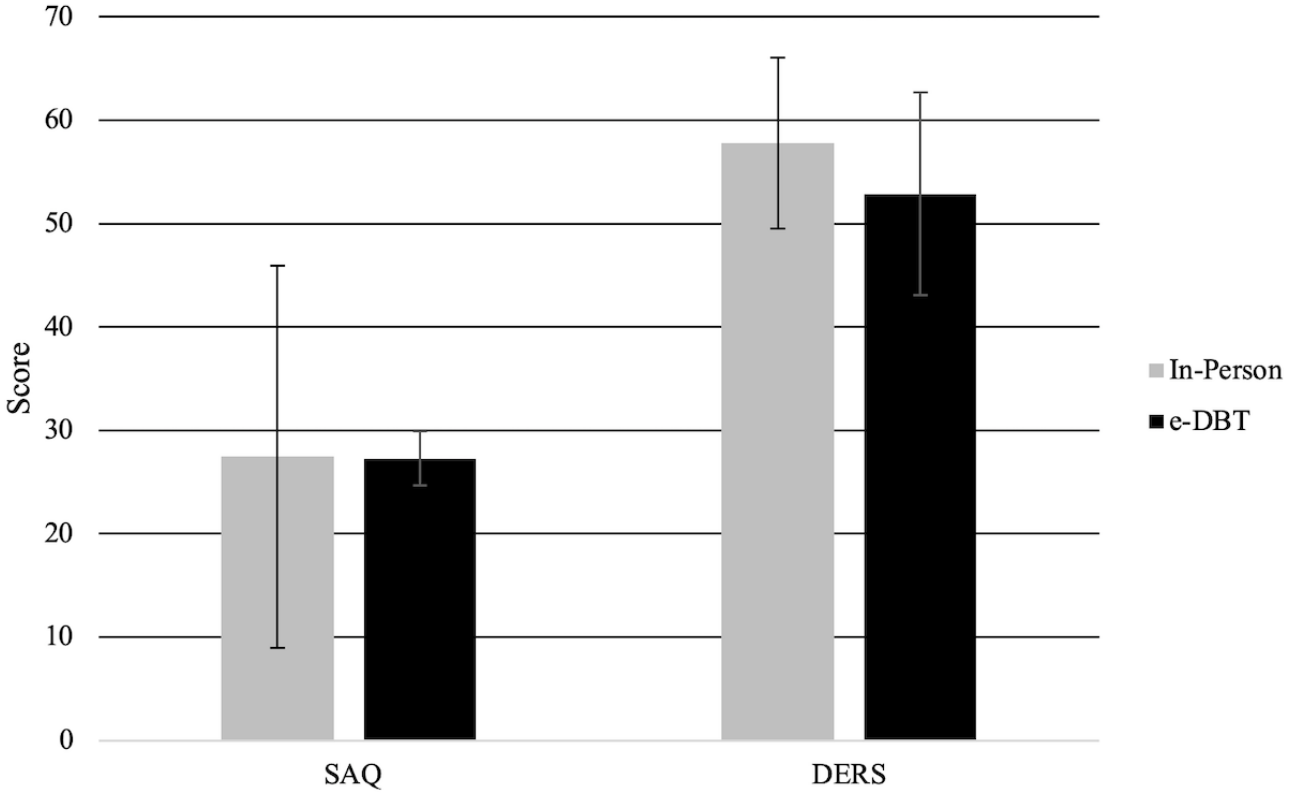


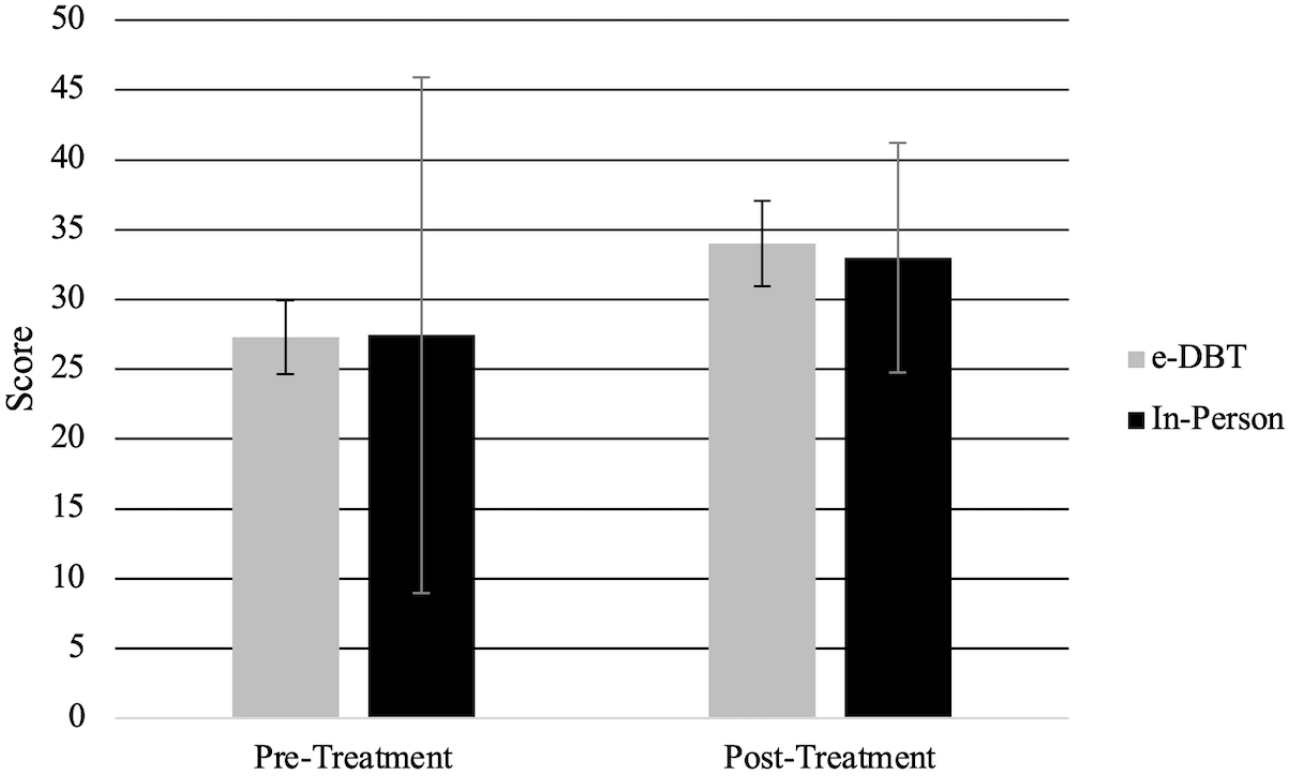
Figure 2. Prior to treatment, there were no significant differences between the 2 groups in Self-Assessment Questionnaire (SAQ) scores and Difficulties in Emotion Regulation Scale (DERS) scores. e-DBT: electronically delivered dialectical behavioral therapy.



SAQ Scores

SAQ scores (Figure 3) were significantly different between pre- and posttreatment ($F_{1,92}$=130.23, P<.001), with pretreatment scores (mean 27.57, SD 7.25) being significantly lower than posttreatment scores (mean 33.28, SD 7.66). Within the groups, significant increases in SAQ scores from pre- to posttreatment in the e-DBT group ($F_{1,92}$=69.32, P<.001) and in-person group ($F_{1,92}$=60.97, P<.001) were observed. There was no significant difference between the groups between pre- and posttreatment for SAQ scores ($F_{1,92}$=0.05, P=.83).

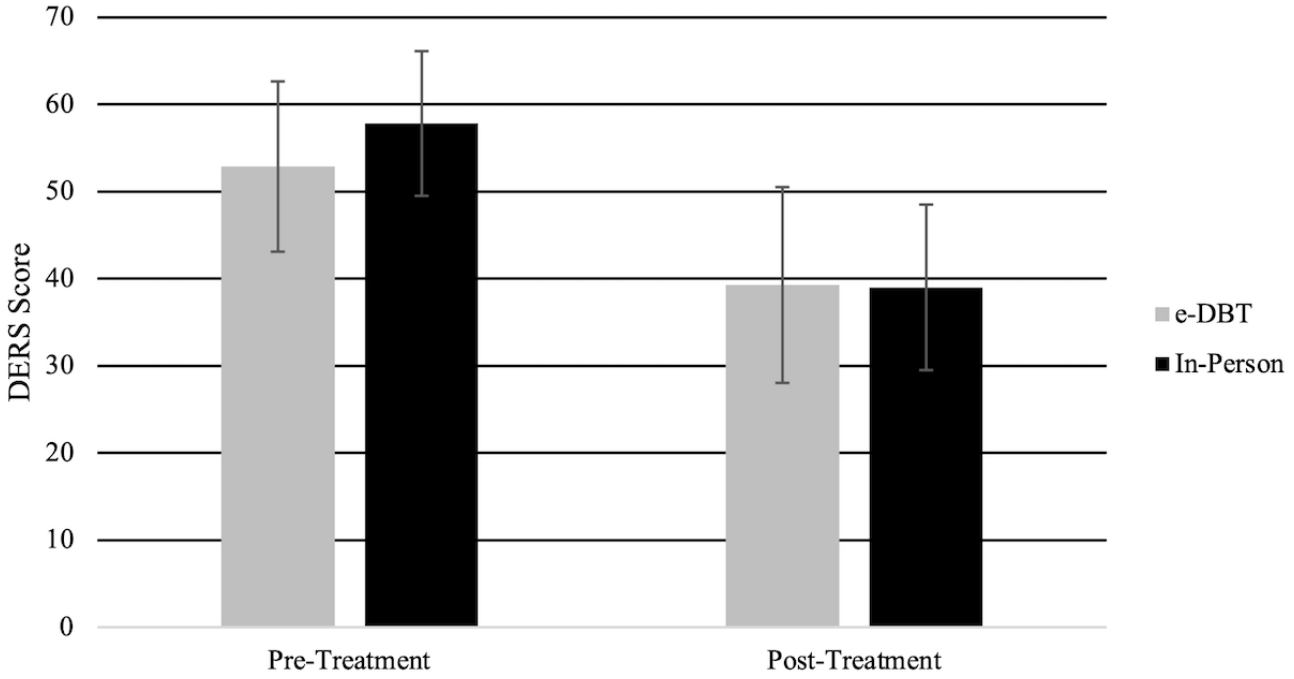
Figure 3. Self-Assessment Questionnaire scores at baseline (pre) and 15 weeks (post). e-DBT: electronically delivered dialectical behavioral therapy.



DERS Scores

DERS scores (Figure 4) were significantly different between pre- and posttreatment ($F_{1,91}$=85.90, P<.001), with pretreatment scores (mean 54.68, SD 22.80) being significantly higher than posttreatment scores (mean 38.78, SD 23.78). Within the groups, there were significant decreases in DERS scores between pre- and posttreatment in the e-DBT group ($F_{1,91}$=30.15, P<.001) and the in-person group ($F_{1,91}$=58.18, P<.001). There was no significant difference between the groups between pre- and posttreatment for DERS scores ($F_{1,91}$=0.24, P=.63).

Figure 4. Difficulties in Emotion Regulation Scale (DERS) scores at baseline and 15 weeks. e-DBT: electronically delivered dialectical behavioral therapy.



Nonacceptance Subscale Scores

The nonacceptance subscale scores significantly differed between pre- and posttreatment ($F_{1,91}$=249.86, P<.001), with pretreatment scores (mean 19.70, SD 5.69) being significantly higher than posttreatment scores (mean 13.86, SD 5.26). Within the groups, there were significant decreases in nonacceptance scores between pre- and posttreatment in the e-DBT group ($F_{1,91}$=123.98, P<.001) and the in-person group ($F_{1,91}$=125.90, P<.001). There was no significant difference between the groups between pre- and posttreatment ($F_{1,91}$=.11, P=.74).

Goals Subscale Scores

The goals subscale scores were significantly different between pre- and posttreatment ($F_{1,92}$=117.10, P<.001), with pretreatment scores (mean 14.15, SD 4.54) being significantly higher than posttreatment scores (mean 10.97, SD 4.71). Within the groups, there were significant decreases in goals scores between pre- and posttreatment in the e-DBT group ($F_{1,92}$=48.96, P<.001) and the in-person group ($F_{1,92}$=69.25, P<.001), with no significant difference between the groups between pre- and posttreatment ($F_{1,92}$=.001, P=.97).

Impulse Subscale Scores

The impulse subscale scores were significantly different between pre- and posttreatment ($F_{1,79}$=131.33, P<.001), with pretreatment scores (mean 12.78, SD 7.74) being significantly higher than posttreatment scores (mean 8.36, SD 5.08). Within the groups, there were significant decreases in impulse scores between pre- and posttreatment in the e-DBT group ($F_{1,79}$=69.01, P<.001) and the in-person group ($F_{1,79}$=63.38, P<.001), with no significant difference between the groups between pre- and posttreatment ($F_{1,79}$=0.005, P=.94).

Awareness Subscale Scores

The awareness subscale scores were significantly different between pre- and posttreatment ($F_{1,76}$=27.31, P<.001), with pretreatment scores (mean –17.22, SD 5.80) being significantly higher than posttreatment scores (mean –19.69, SD 6.92). Within the in-person group, there were significant differences between pre- and posttreatment scores ($F_{1,76}$=32.20, P<.001); however, within the e-DBT group, there was no difference between pre- and posttreatment scores, ($F_{1,76}$=1.84, P=.18). There was no significant difference between the groups between pre- and posttreatment ($F_{1,76}$=.71, P=.40).

Strategies Subscale Scores

The strategies subscale scores were significantly different between pre- and posttreatment ($F_{1,91}$=171.73, P<.001), with pretreatment scores (mean 21.73, SD 7.12) significantly higher than posttreatment scores (mean 15.62, SD 6.90). Within the groups, there were significant decreases in strategies scores between pre- and posttreatment in the e-DBT group ($F_{1,91}$=79.10, P<.001) and the in-person group ($F_{1,91}$=93.00, P<.001), with no significant difference between the groups between pre- and posttreatment ($F_{1,91}$=0.013, P=.91).

Clarity Subscale Scores

The clarity subscale scores were significantly different between pre- and posttreatment ($F_{1,75}$=121.03, P<.001), with pretreatment scores (mean 2.69, SD 4.00) significantly higher than posttreatment scores (mean –1.18, SD 3.65). Within the groups, there were significant decreases in clarity scores between pre- and posttreatment in the e-DBT group ($F_{1,91}$=78.85, P<.001) and the in-person group ($F_{1,91}$=47.65, P<.001), with no significant difference between the groups between pre- and posttreatment ($F_{1,75}$=0.005, P=.94).

Discussion

General

DBT is a form of psychotherapy that has been proven to be an efficacious treatment modality for addressing various mental health disorders in several controlled research studies [7]. DBT is particularly effective in reducing the incidence of suicidal and self-injurious behaviors and the frequency of acute hospitalizations in individuals diagnosed with borderline personality disorder [20].

Significance and Impact

Many individuals with borderline personality disorder are resistant to taking part in in-person group psychotherapy, a core aspect of DBT [20,21]. Additionally, there are many other psychological, social, geographical, and systemic barriers to utilizing DBT as a treatment for mental health disorders [9]. The demand for DBT often exceeds the resources, leaving many individuals with serious and life-threatening (in some cases) problems, on waitlists for evidence-based care. Therefore, it is an unequivocal public health need to overcome these barriers through alternative methods of care delivery. With internet use increasing globally, offering internet-based DBT skill-building groups through email (e-DBT) could be a viable treatment option that could help the health care system meet the demand for therapy. The ability to reduce treatment costs while offering comparable quality of care with more efficient utilization of medical personnel can be significant to the health care system. Among other benefits, e-DBT would allow for greater treatment accessibility to participants, as well as being more time-efficient for clinicians without having to sacrifice the quality of care. Moreover, e-DBT would allow for relatively simple modifications in the future when addressing language and cultural barriers to therapy. Additionally, e-DBT can provide a much-needed service to individuals located in geographically isolated areas.

Our results suggest that an e-DBT skill-training program delivered via email could be a viable treatment delivery modality for addressing symptoms in individuals diagnosed with borderline personality disorder. There were no significant differences observed in SAQ or DERS scores between the e-DBT and in-person groups both pre- and posttreatment (P<.001). This suggests that e-DBT could provide comparable results to those provided by in-person therapy, allowing a more accessible version of the treatment, without sacrificing the quality of care. Additionally, both the SAQ and DERS scores significantly improved in both the e-DBT and in-person groups (P<.001). These results suggest that e-DBT could be an effective alternative to in-person therapy for individuals with borderline personality disorder.

Although there was no significant difference observed between the groups in terms of the number of participants who completed the program (P<.001), participants who prematurely terminated their involvement in the e-DBT program took part in more sessions than those who prematurely terminated participation in the in-person group. This could indicate that e-DBT offers a higher treatment adherence in individuals with borderline personality disorder.

Limitations

Despite the strengths of this study, there are some limitations. The study did not assess the long-term efficacy of the treatment; future research should investigate this by implementing a follow-up component.

Additionally, among the participants who selected the e-DBT and in-person groups, only 44% (23/52) and 49% (27/55), respectively, completed the program. The large number of dropouts could affect the result of the study; however, we believe that the lack of adherence with treatment could be due to the nature of borderline personality disorder. For individuals with borderline personality disorder, the dropout rates in a DBT outpatient group are typically quite high, often peaking between 24% and 58% and are attributed to a younger age, higher levels of baseline distress, and a higher level of baseline nonacceptance of emotional responses [22]; therefore, the dropout rates in our study are not unusual

Future Direction

Although the therapy at the Personality Disorder Services at Queen's University, Kingston, Ontario is offered in 3 phases—DBT-informed skill-building group (Managing Powerful Emotions), psychotherapy groups, Chrysalis Program—in this study, electronic delivery was only offered for the first phase. Future research should evaluate the efficacy of electronic delivery of all 3 phases.

Future research is necessary to address the abovementioned limitations and to provide further support for the efficacy of an e-DBT program. Although the results of our study suggest that email is a viable method for delivering DBT skill-building groups, a randomized controlled study should be conducted to compare the efficacy of in-person DBT with e-DBT for borderline personality disorder treatment. A control group should be utilized to examine its efficacy in comparison to other delivery methods. Moreover, a benefit to randomization would be that individuals with differing technology comfort levels would be more evenly dispersed. Future research should also implement a follow-up period to ensure that e-DBT has long-term efficacy.

Conclusions

Notwithstanding its limitations, our study's findings have significant practical implications. This study provides evidence that DBT delivered via email can be effective in reducing the severity of symptoms associated with borderline personality disorder. These findings concurrently add to literature on the

efficacy of internet-based interventions for DBT, with more work needing to be done [23]. This innovative modality has the potential to increase the accessibility of mental health services for a large group of individuals who could benefit from these resources. This simple, innovative, and user-friendly way to deliver DBT can be used to deal with barriers to treatment such as lack of resources, work or school commitments, transportation limitations, geographical isolation, the stigma associated with mental health treatment, and the high costs of software development. This treatment modality can be particularly beneficial for those comfortable with technology who may be concerned with the stigma associated with attending in-person DBT or group treatments by allowing treatment to be completed at any time and location. The DBT delivered via email shows promise as a new treatment delivery modality that can provide increased accessibility while offering improvements in symptomology that are comparable to of in-person DBT for individuals with borderline personality disorder.

Acknowledgments

This study was funded by the Psychiatry Department, Queen's University in Kingston, Ontario, Canada (internal grant 376051). We would like to thank Karen Gagnon for all her hard work during this study and dedicate this paper to her memory.

Authors' Contributions

NA and MR conceived and designed the study and acquired, analyzed, and interpreted data. NA MR, and CS drafted the manuscript. CS structured and edited the manuscript. All authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

NA cofounded the care delivery platform in use (OPTT) and has an ownership stake in OPTT Inc.

References

1. Sanislow CA, McGlashan TH. Treatment outcome of personality disorders. Can J Psychiatry 1998 Apr;43(3):237-250. [doi: 10.1177/070674379804300302] [Medline: 9561312]
2. Andreoli A, Gressot G, Aapro N, Tricot L, Gognalons MY. Personality disorders as a predictor of outcome. J Pers Disord 1989 Dec;3(4):307-320. [doi: 10.1521/pedi.1989.3.4.307]
3. Reich JH, Green AI. Effect of personality disorders on outcome of treatment. J Nerv Ment Dis 1991 Feb;179(2):74-82. [doi: 10.1097/00005053-199102000-00003] [Medline: 1990074]
4. Reich JH, Vasile RG. Effect of personality disorders on the treatment outcome of axis I conditions: an update. J Nerv Ment Dis 1993 Aug;181(8):475-484. [doi: 10.1097/00005053-199308000-00002] [Medline: 8103074]
5. Chapman J, Jamil RT, Fleisher C. Borderline Personality Disorder. Treasure Island: StatPearls Publishing; 2020.
6. Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF. Evaluating three treatments for borderline personality disorder: a multiwave study. Am J Psychiatry 2007 Jun;164(6):922-928. [doi: 10.1176/ajp.2007.164.6.922] [Medline: 17541052]
7. Linehan M. Cognitive–Behavioral Treatment of Borderline Personality Disorder. New York: Guilford Press; 1993:1-1.
8. McMain SF, Guimond T, Streiner DL, Cardish RJ, Links PS. Dialectical behavior therapy compared with general psychiatric management for borderline personality disorder: clinical outcomes and functioning over a 2-year follow-up. Am J Psychiatry 2012 Jun;169(6):650-661. [doi: 10.1176/appi.ajp.2012.11091416] [Medline: 22581157]
9. Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers A, et al. Mind the gap: improving the dissemination of CBT. Behav Res Ther 2009 Nov;47(11):902-909. [doi: 10.1016/j.brat.2009.07.003] [Medline: 19664756]
10. Townsend L, Gearing RE, Polyanskaya O. Influence of health beliefs and stigma on choosing internet support groups over formal mental health services. Psychiatr Serv 2012 Apr;63(4):370-376. [doi: 10.1176/appi.ps.201100196] [Medline: 22337006]
11. Chen C, Wang C. Internet as indispensable everywhere: the introduction to the advances in internet technologies and applications special issue. J Comput 2013 Jul 01;8(7):1-1. [doi: 10.4304/jcp.8.7.1633-1634]
12. Alavi N, Hirji A. The efficacy of PowerPoint-based CBT delivered through email: breaking the barriers to treatment for generalized anxiety disorder. J Psychiatr Pract 2020 Mar;26(2):89-100. [doi: 10.1097/PRA.0000000000000455] [Medline: 32134882]
13. Ruwaard J, Broeksteeg J, Schrieken B, Emmelkamp P, Lange A. Web-based therapist-assisted cognitive behavioral treatment of panic symptoms: a randomized controlled trial with a three-year follow-up. J Anxiety Disord 2010 May;24(4):387-396. [doi: 10.1016/j.janxdis.2010.01.010]
14. Ruwaard J, Lange A, Schrieken B, Dolan CV, Emmelkamp P. The effectiveness of online cognitive behavioral treatment in routine clinical practice. PLoS One 2012;7(7):e40089 [FREE Full text] [doi: 10.1371/journal.pone.0040089] [Medline: 22792217]
15. 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] [Medline: 20183695]

16. Crabb RM, Cavanagh K, Proudfoot J, Learmonth D, Rafie S, Weingardt KR. Is computerized cognitive-behavioural therapy a treatment option for depression in late-life? a systematic review. Br J Clin Psychol 2012 Nov;51(4):459-464. [doi: 10.1111/j.2044-8260.2012.02038.x] [Medline: 23078214]
17. Rivera M. Rethinking mental health and disorder: feminist perspectives. Arch Womens Ment Health 2003 Aug 1;6(3):219-219. [doi: 10.1007/s00737-003-0016-6]
18. Rivera M, Darke JL. Integrating empirically supported therapies for treating personality disorders: a synthesis of psychodynamic and cognitive-behavioral group treatments. Int J Group Psychother 2012 Oct;62(4):500-529. [doi: 10.1521/ijgp.2012.62.4.500] [Medline: 22974149]
19. Weiss NH, Gratz KL, Lavender JM. Factor structure and initial validation of a multidimensional measure of difficulties in the regulation of positive emotions: the DERS-Positive. Behav Modif 2015 May;39(3):431-453 [FREE Full text] [doi: 10.1177/0145445514566504] [Medline: 25576185]
20. Alper G, Peterson S. Dialectical behavior therapy for patients with borderline personality disorder. J Psychosoc Nurs Ment Health Serv 2001 Oct;39(10):38-45. [Medline: 11697073]
21. Verheul R, Van Den Bosch LMC, Koeter MWJ, De Ridder MAJ, Stijnen T, Van Den Brink W. Dialectical behaviour therapy for women with borderline personality disorder: 12-month, randomised clinical trial in The Netherlands. Br J Psychiatry 2003 Feb;182:135-140. [doi: 10.1192/bjp.182.2.135] [Medline: 12562741]
22. Landes SJ, Chalker SA, Comtois KA. Predicting dropout in outpatient dialectical behavior therapy with patients with borderline personality disorder receiving psychiatric disability. Borderline Personal Disord Emot Dysregul 2016;3(1):9 [FREE Full text] [doi: 10.1186/s40479-016-0043-3] [Medline: 27588205]
23. Frías Á, Solves L, Navarro S, Palma C, Farriols N, Aliaga F, et al. Technology-based psychosocial interventions for people with borderline personality disorder: a scoping review of the literature. Psychopathology 2020;53(5-6):254-263 [FREE Full text] [doi: 10.1159/000511349] [Medline: 33166964]

Abbreviations

DBT: dialectical behavioral therapy
DERS: Difficulties in Emotion Regulation Scale
e-CBT: electronically delivered cognitive behavioral therapy
e-DBT: electronically delivered dialectical behavioral therapy
SAQ: Self-Assessment Questionnaire

Edited by J Torous; submitted 21.01.21; peer-reviewed by A Anastasiou, S Pitoglou; comments to author 03.03.21; revised version received 04.03.21; accepted 09.03.21; published 30.04.21.

Please cite as:
Alavi N, Stephenson C, Rivera M
Effectiveness of Delivering Dialectical Behavioral Therapy Techniques by Email in Patients With Borderline Personality Disorder: Nonrandomized Controlled Trial
JMIR Ment Health 2021;8(4):e27308
URL: https://mental.jmir.org/2021/4/e27308
doi:10.2196/27308
PMID:33835936

Original Paper

The Effect of Meditation and Physical Activity on the Mental Health Impact of COVID-19–Related Stress and Attention to News Among Mobile App Users in the United States: Cross-sectional Survey

Jennifer Green[1], PhD; Jennifer Huberty[1], PhD; Megan Puzia[2], MS; Chad Stecher[1], PhD

[1]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

[2]Behavioral Research and Analytics, LLC, Salt Lake City, UT, United States

Corresponding Author:
Jennifer Green, PhD
College of Health Solutions
Arizona State University
500 N. 3rd St.
Phoenix, AZ, 85004
United States
Phone: 1 602 827 2314
Email: jeni.green@asu.edu

Abstract

Background: The COVID-19 pandemic has been declared an international public health emergency, and it may have long-lasting effects on people's mental health. There is a need to identify effective health behaviors to mitigate the negative mental health impact of COVID-19.

Objective: The objectives of this study were to (1) examine the regional differences in mental health and COVID-19–related worry, attention to news, and stress, in light of the state-level prevalence of COVID-19 cases; (2) estimate the associations between mental health and COVID-19–related worry, attention to news, and stress and health behavior engagement (ie, physical activity, mindfulness meditation); and (3) explore the mediating effect of health behavior engagement on the associations between mental health and COVID-19–related worry, attention to news, and stress.

Methods: A cross-sectional survey was distributed to a sample of US adult paying subscribers to the Calm app (data were collected from April 22 to June 3, 2020). The survey assessed COVID-19–related worry, attention to news, and stress; health behavior engagement; and mental health (ie, perceived stress, posttraumatic stress disorder, and anxiety and depression). Statistical analyses were performed using R software. Differences in COVID-19–related worry, attention to news, and stress and mental health by location were assessed using *t* tests and chi-square tests. Logistic and ordinary least squares models were used to regress mental health and health behavior on COVID-19–related worry, attention to news, and stress; moreover, causal mediation analysis was used to estimate the significance of the mediation effects.

Results: The median age of the respondents (N=8392) was 47 years (SD 13.8). Participants in the Mid-Atlantic region (New Jersey, New York, and Pennsylvania) reported higher levels of stress, more severe depression symptoms, greater worry about COVID-19, paying more attention to COVID-19–related news, and more stress related to social distancing recommendations than participants living in other regions. The association between worry about COVID-19 and perceived stress was significantly mediated by changes in physical activity (P<.001), strength of meditation habit (P<.001), and stopping meditation (P=.046). The association between worry about COVID-19 and posttraumatic stress disorder symptoms was significantly mediated by changes in physical activity (P<.001) and strength of meditation habit (P<.001).

Conclusions: Our findings describe the mental health impact of COVID-19 and outline how continued participation in health behaviors such as physical activity and mindfulness meditation reduce worsening of mental health due to the COVID-19 pandemic. These data have important implications for public health agencies and health organizations to promote the maintenance of health habits to reduce the residual mental health burden of the COVID-19 pandemic.

(JMIR Ment Health 2021;8(4):e28479) doi:10.2196/28479

KEYWORDS

coronavirus; health behavior; mindfulness meditation; mHealth; COVID-19; mental health

Introduction

In January 2020, the World Health Organization declared COVID-19 an international public health emergency [1], and the negative mental health effects of the ongoing COVID-19 pandemic are expected to be a significant, long-lasting global health problem [2,3]. In an April 2020 review of the existing literature on COVID-19 and mental health, moderate to severe levels of depressive symptoms and anxiety were reported in 16%-28% of the general population and medical staff in response to the COVID-19 pandemic [4]. Additionally, fear and worry about COVID-19 are common [5], with many people citing worries related to personal infection or the infection of family members [6,7], an overrun health care system, financial losses without expectation of recovery soon [8,9], and long-lasting isolation and movement restrictions [10]. Previously reported data have suggested regional differences across the United States in COVID-19–related fear and mental health (ie, anxiety and depressive symptoms), with greater symptoms in regions with higher confirmed cases, namely the Northeast New England, Northeast Mid-Atlantic, South-South Atlantic, and West Pacific regions (survey data collected March 23, 2020) [5]. Based on the known mental health effects of the COVID-19 pandemic, there is a clear need for strategies to help individuals better cope with the pandemic and mitigate its potentially long-lasting mental health consequences. Additionally, there is a need to better understand how resources for mental health should be allocated as the prevalence of COVID-19 infections changes regionally over time.

Although mental health professionals are often among the first line of treatment for poor mental health, digital approaches, including mobile health (mHealth) technologies may provide a way to more widely disseminate treatment information and enable individuals to self-manage their mental health from the safety of their own home. In a recent survey (N=2198) by the Academy of Medical Services, many respondents were concerned about how they would access mental health support, as many previously available in-person services had been discontinued as a result of the COVID-19 pandemic [11]. Stakeholders were also concerned with the capacity to handle the increased demand for mental health services and the lack of emphasis on mental health compared to treatment of COVID-19 and its physical health impacts [11]. Importantly, self-management strategies, including digital approaches to improve mental health, have become an area of interest for policy makers, as many individuals fail (or are unable) to participate due to the pandemic (eg, facility closures, reduced client load, social distancing) [12]. Self-management strategies may also empower individuals by enabling them to take a more active role in their health care (ie, recognizing and managing their own health problems) [13]. Additionally, self-management strategies are cost-effective and can be used as a preventative tool (rather than prescriptive or treatment-focused) that may mitigate the development of more debilitating mental health issues [12]. Common and evidence-based self-management strategies for improving mental health are physical activity and mindfulness meditation [14-16], both of which can still be maintained during social distancing policies and stay-at-home orders and can be adapted and delivered digitally. However, the extent to which individuals have maintained their participation in these self-management strategies during the COVID-19 pandemic is unknown.

Physical activity has been widely adopted as a beneficial way to self-manage mental health and may attenuate the mental health decline resulting from COVID-19 [17,18]. Physical activity has been shown to be as effective as antidepressants in decreasing stress, improving mood, and enhancing self-esteem [17,18]. Despite the known benefits of physical activity, currently, more than 60% of US adults do not engage in the recommended amount of physical activity (ie, 150 minutes of moderate to vigorous physical activity per week), and 25% are not active at all (ie, sedentary) [19]. Social distancing, quarantine or social isolation, and closure of public spaces due to COVID-19 are likely to increase these rates of physical inactivity [20,21], and the degree to which adults change the type or duration of their physical activity during COVID-19 is also unknown. Thus, there is a need to determine how self-management strategies such as physical activity have changed as a result of the pandemic, and specifically if reductions in physical activity are associated with worsening mental health status.

Mindfulness meditation, another self-management strategy, has also been evidenced to improve mental health, and maintenance (or habituation) of this behavior, particularly during the COVID-19 pandemic, may reduce worsening of mental health over time. Evidence suggests that mindfulness meditation reduces stress, improves mood (eg, symptoms of anxiety and depression), and enhances well-being [14,22]. Although an optimal amount of mindfulness meditation has not been established, evidence suggests that its benefits accrue with greater frequency of practice [23]. The habituation of mindfulness meditation practice (ie, behavioral automaticity) [24,25] has yet to be examined in the literature [26] but may have important implications for whether individuals maintain their practice during the COVID-19 pandemic. Those with stronger meditation habits may be more likely to continue meditation practice even when daily life is disrupted, as habits are known to persist when motivation declines and other distractions are present [25]. Mindfulness meditation delivered via mHealth may be another useful way to help people maintain their practice, especially when many public facilities are closed or limit participation [27-29], and popular meditation apps such as Calm or Headspace have shown promise to reduce stress and improve health [27,30-34]. Interestingly, sales of the Calm app in April 2020 were 62% higher than in February 2020 and 32% higher than in March 2020 (unpublished sales data provided by an internal Calm team), indicating that more people are accessing this type of self-management strategy in response to COVID-19. Although mHealth meditation apps have promise to help people self-manage their mental health, there is a lack

of understanding about how continued participation in meditation is being impacted during the COVID-19 pandemic and whether meditation habits are associated with improved mental health.

Given the benefits of physical activity and mindfulness meditation on mental health, the purpose of this cross-sectional survey (data collected between April 22 and June 3, 2020) was to (1) examine the differences in COVID-19–related worry, attention to news, and stress from social distancing and mental health (ie, stress, posttraumatic stress disorder [PTSD], depression, and anxiety) by region of the United States, (2) explore the associations between COVID-19–related worry, attention to news, and stress and health behavior engagement (ie, strength of meditation habit, and changes in mindfulness meditation and physical activity) and mental health, and (3) estimate the mediating effect of health behavior engagement on the associations between mental health and COVID-19–related worry, attention to news, and stress. We hypothesized that (1) there would be differences in COVID-19–related worry, attention to news, and stress and mental health in states with higher prevalence of COVID-19, (2) greater COVID-19–related worry, attention to news, and stress would be associated with lower health behavior engagement and greater levels of mental health, and (3) health behavior engagement would mediate the associations between mental health and COVID-19–related worry, attention to news, and stress.

Methods

Ethics Approval

The Institutional Review Board at Arizona State University (STUDY00011867) approved the study. All participants provided electronic consent before participating in the study. The data sets generated and analyzed during the study are available at OSF [35].

Study Design

This survey is part of a descriptive, national longitudinal study using a nonrandom convenience sample of adult paying subscribers to Calm, a mindfulness meditation mobile app. Participants in this study initially completed a cross-sectional baseline survey called the “COVID-19 Health and Well-being Survey” and agreed to complete four follow-up surveys over the subsequent 12 months. The data presented in this paper were obtained from the baseline survey administered from April 22 to June 3, 2020.

Participant Recruitment and Selection

Emails inviting Calm subscribers to participate in the study were sent on April 22, April 29, and May 6, 2020. Subscribers were eligible if they had opened an email from Calm at least once in the last 90 days and had used Calm at least once in the last 90 days, were aged 18 years or older, could read and understand English, and resided in the United States.

Procedures

Interested individuals were directed to a Qualtrics eligibility screener (~1 minute to complete). Once eligibility was determined, participants completed an electronic informed consent form and the baseline survey. There were no incentives for participation in the first wave (baseline) of the study; however, incentives were offered for continued participation in subsequent waves of this study (ie, random draws for 20 US $50 Amazon gift cards at months 2-4 and 50 US $50 Amazon gift cards at month 12).

COVID-19 Health and Well-being Survey

The baseline survey was also administered using Qualtrics and included both investigator-developed and validated questionnaires. The investigator-developed portion of the survey included a total of 15 questions related to worry regarding COVID-19, attention to news, and stress from social distancing and 20 questions related to health behavior engagement (ie, strength of meditation habit and changes in mindfulness meditation and physical activity; see Textbox 1 for the questions used in these analyses). The COVID-19–related questions were adapted from a US Centers for Disease Control and Prevention (CDC) questionnaire for other infectious diseases (eg, the Severe Acute Respiratory Syndrome [SARS] Psychosocial Research Consortium survey) [36,37]. The validated survey components assessed meditation habit strength, perceived stress, PTSD, and anxiety and depression. Demographic information was collected at the end of the survey. Specifically, zip codes were used to determine the participants’ state of residence [38], and the states were categorized into regions according to the US Census Bureau classification system [39]. States were also designated as having a high or low prevalence of COVID-19 based on data compiled by the Center for Systems Science and Engineering at Johns Hopkins University describing state-level COVID-19 cases and deaths per 100,000 at the time of the initial survey distribution (April 22, 2020) [40].

Textbox 1. Investigator-developed survey questions used in the analyses.

Prior to COVID did you meditate with Calm? [Y/N]

Do you currently meditate with Calm? [Y/N]

To what extent has COVID changed your meditation practice?

- I meditate much more (5)
- I meditate a little more (4)
- I meditate about the same amount (3)
- I meditate a little less (2)
- I meditate much less (1)
- I no longer meditate (0)

Prior to COVID, on average, how often did you engage in physical activity/exercise? [0-7 days/week]

On average, how often do you currently engage in physical activity/exercise? [0-7 days/week]

To what extent has COVID-19 changed your amount of physical activity/exercise?

- I exercise much more (5)
- I exercise a little more (4)
- I exercise about the same amount (3)
- I exercise a little less (2)
- I exercise much less (1)
- I no longer exercise (0)

Have recommendations for socially distancing caused stress for you?

- Not at all (1)
- A little (2)
- Somewhat (3)
- A lot (4)

How worried are you about…

… personally getting coronavirus?

… a family member getting coronavirus?

… the spread of coronavirus in your area?

- Not at all worried (1)
- A little worried (2)
- Somewhat worried (3)
- A good bit worried (4)
- Very worried (5)

How would you rate your attentiveness to information about ongoing changes and updates regarding the coronavirus?

- Not at all paying attention to (1)
- Somewhat paying attention to (2)
- Moderately paying attention to (3)
- Quite a bit paying attention to (4)
- Very much paying attention to (5)

Measures

COVID-19–Related Worry, Attention to News, and Stress From Social Distancing

Participants were asked how worried they were about personally contracting COVID-19, a family member contracting COVID-19, and the spread of COVID-19 in their area (see Textbox 1). Worry about COVID-19 was operationalized as the sum of responses to these three questions. Scores ranged from 3-15, where higher scores indicate greater levels of worry about COVID-19. Participants were also asked how they would rate their attentiveness to information about ongoing changes and news regarding COVID-19 (see Textbox 1). Scores at or below the median were used to identify individuals with low attention to COVID-19, while scores above the median were categorized as high attention to COVID-19. Finally, the participants were asked if recommendations for socially distancing had caused them stress (see Textbox 1). Scores at or below the median were used to identify participants with low stress; scores above the median indicated high stress.

Health Behavior Engagement

Mindfulness Meditation Practice

Participants were asked about their meditation practice using Calm prior to the COVID-19 pandemic as well as their current use of Calm. If participants indicated that they meditated using Calm prior to the COVID-19 pandemic, they were asked to what extent COVID-19 had changed their meditation practice (see Textbox 1). Participants who indicated they no longer participated in meditation were categorized as "stopped meditation," while all other participants were considered as "continuing meditation."

Physical Activity Behavior

Participants were asked to select how many days per week (scale of 0-7) they participated in physical activity prior to the COVID-19 pandemic as well as their current participation. Participants were also asked to what extent COVID-19 had changed the frequency or duration of their physical activity (see Textbox 1). Changes in physical activity were calculated as the difference between the participants' reported number of days of physical activity currently and prior to the COVID-19 pandemic.

Validated Surveys

Self-Report Habit Index

The Self-Report Habit Index (SRHI) includes 12 items reflecting on three proposed characteristics of habit (ie, automaticity, frequency, and relevance to self-identity). Response options range from 1, strongly disagree, to 5, strongly agree. The SRHI is a reliable and valid measure that has demonstrated Cronbach α values of .89-.92. The items are summed to produce a total score, with higher values indicating stronger habits [41]. For the current study, only the 4-item automaticity subscale of the SRHI was used in the analyses, and it produced a Cronbach α of .88.

Perceived Stress Scale

The Perceived Stress Scale (PSS) includes 10 items that measure the degree of self-appraised stress in one's life within the past month [42,43]. The response items are rated on a 5-point Likert scale from 0, never, to 4, very often. The items are summed to produce a total score from 0-40, with higher scores indicating higher levels of perceived stress. The PSS is a reliable and valid measure that has demonstrated good internal consistency (Cronbach α=.74-.91) [44]. For the current study, the Cronbach α is .89.

Impact of Events Scale-6

The Impact of Events Scale-6 (IES-6) is a 6-item abbreviated version of the Impact of Events Scale-Revised (22 items) that assesses PTSD symptoms over the past seven days [45]. The response items are rated on a 5-point Likert scale from 0, not at all, to 4, extremely. The score is calculated as a mean of the six items. Scores range from 0-5, with a binary cutoff score of 1.75 indicating clinically important PTSD symptoms. The IES-6 is a valid and reliable measure with excellent internal consistency (α=.91). For the current study, the Cronbach α is .86.

Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) is a 14-item scale measuring levels of anxiety and depression [46]. The anxiety subscale (HADS-A) and the depression subscale (HADS-D) each comprise 7 items. Response items are rated on a 4-point Likert scale from 0-3. The items are summed to produce a total score from 0-21 on each subscale. Scores from 0-7 are considered normal, scores from 8-10 are considered borderline abnormal, and scores from 11-21 are considered abnormal. The HADS is a valid and reliable tool, with internal consistencies reported to be as high as α=0.93 and α=0.90 for the HADS-A and HADS-D subscales, respectively. For the current study, the Cronbach α=.85 and .80 for the HADS-A and HADS-D, respectively.

Statistical Analysis

All statistical analyses were performed using R software, version 4.0.0 (R Project) [47]. Descriptive statistics were used to characterize the sample's demographic characteristics; health; presence of chronic conditions; mental health; and COVID-19–related worry, attention to news, and stress. Differences in mental health and COVID-19–related worry, attention to news, and stress by location were assessed using *t* tests and chi-square tests. Logistic and ordinary least squares models were used to regress mental health and health behavior on COVID-19–related worry, attention to news, and stress, and causal mediation analysis from the "mediation" package in R was used to estimate the significance of the mediation effects. Specifically, the standard errors for the mediation effects were calculated from 1000 bootstrapped samples for each mediation regression. Demographic, health, and location variables were included as covariates in all regression analyses. A *P* value of <.05 was considered statistically significant.

Results

Sample Characteristics

The sample (N=8392) was primarily White, non-Hispanic, and female (see Table 1). The majority of participants had a bachelor's or graduate-level degree, were employed, and had an annual household income exceeding US $100,000. Approximately one-third of the participants (2603/7335, 35.49) reported having a least one medical condition associated with increased risk of severe illness from COVID-19; however, more than 80% (5976/7317, 81.67%) perceived themselves to be in good overall health.

The self-reported changes in preventative health behaviors are shown in Table 2; most participants were found to have increased or maintained their physical activity and meditation habits during the initial period of the COVID-19 pandemic. There was a significant correlation between strength of meditation habit and changes in meditation during the COVID-19 pandemic; participants with the strongest habits were the most likely to increase or maintain their meditation practices (r=0.37, P<.001).

The mental health characteristics of the sample at the time of the baseline survey are presented in Table 3. Compared to participants living in other parts of the country, participants living in the Mid-Atlantic region (ie, New Jersey, New York, or Pennsylvania) reported higher levels of stress and more severe depression symptoms. Participants living in the South Atlantic region (ie, Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, or District of Columbia) reported less severe depressive symptoms than participants in other regions. Participants living in states with a high prevalence of COVID-19 cases (ie, California, Colorado, Illinois, Massachusetts, New Jersey, New York, or Washington) reported more severe PTSD symptoms than participants living in states where COVID-19 was less prevalent.

Table 4 presents the participants' reports regarding their worries about COVID-19 (ie, contracting it themselves, a family member contracting it, the spread in their area), their attentiveness to news and updates regarding COVID-19, and their stress due to COVID-19–related social distancing recommendations. Participants in the Mid-Atlantic region were more worried about COVID-19, paid more attention to COVID-19 news and updates, and reported more stress related to social distancing recommendations than participants living in other regions. Conversely, participants living in the South Atlantic region reported less stress from social distancing recommendations than participants in other regions. Participants living in states where COVID-19 was prevalent had more COVID-19–related worry, paid more attention to COVID-19 news and updates, and experienced more stress from social distancing recommendations.

The primary sources participants used to acquire information about COVID-19 were news media (eg, newspapers, web-based newspapers, television news networks; n=5552/7371, 75.32%) and health officials (eg, CDC, World Health Organization, state health officials; n=5212/7371, 75.12%). Approximately one-third of participants reported acquiring information about COVID-19 from social media (n=2951/7371, 40.04%), doctors and medical professionals (n=2721/7371, 36.91%), and friends and family members (n=2543/7371, 34.50%).

Table 1. Demographic and health characteristics of the sample (N=8392).

Characteristic	Value
Age (years), median (SD)	47.0 (13.8)
Gender (n=7303), n (%)	
Female	6129 (83.92)
Male	1147 (15.71)
Other	27 (0.37)
Race (n=7178), n (%)	
White	6586 (91.75)
Black/African American	231 (3.22)
Asian	216 (3.01)
Native American/Alaska Native	83 (1.16)
Native Hawaiian/Pacific Islander	27 (0.38)
Other	195 (2.72)
Ethnicity (n=6774), n (%)	
Non-Hispanic	6338 (93.56)
Hispanic	436 (6.44)
People in the household, median (SD)	2.0 (1.4)
Income (US $; n=6949), n (%)	
20,000 or less	212 (3.05)
21,000-40,000	402 (5.79)
41,000-60,000	705 (10.15)
61,000-80,000	942 (13.56)
81,000-100,000	1055 (15.18)
More than 100,000	3633 (52.28)
Employment (n=7297), n (%)	
Employed	5084 (69.67)
Retired	1012 (13.87)
Unemployed	477 (6.54)
Homemaker	306 (4.19)
Unable to work	252 (3.45)
Student	166 (2.27)
Education (n=7319), n (%)	
11th grade or less	8 (0.11)
High school or General Educational Development	161 (2.20)
Some college	826 (11.29)
Two-year/technical degree	424 (5.79)
Bachelor's degree	2670 (36.48)
Graduate degree	3230 (44.13)
Region[a] (n=7037), n (%)	
New England	489 (6.95)
Mid-Atlantic	946 (13.44)
East North Central	953 (13.54)
West North Central	459 (6.52)

Characteristic	Value
South Atlantic	1192 (16.94)
East South Central	198 (2.81)
West South Central	522 (7.42)
Mountain West	713 (10.13)
Pacific West	1565 (22.24)
State COVID-19 prevalence[b] (n=7037), n (%)	
High-prevalence state	3062 (43.51)
Low-prevalence state	3975 (56.49)
Health rating (n=7317), n (%)	
Poor	1341 (18.33)
Good	5976 (81.67)
Underlying medical conditions[c] associated with increased risk of severe illness (n=7335), n (%)	
At least one underlying medical condition	2603 (35.49)
No underlying medical conditions	4732 (64.51)

[a]States within each region are based on US Census divisions.

[b]State-level COVID-19 prevalence is based on the number of COVID-19 cases per 100,000 at the time of the first survey distribution (April 22, 2020); high-prevalence states were California, Colorado, Illinois, Massachusetts, New Jersey, New York, and Washington.

[c]The underlying medical conditions associated with increased risk of severe illness from COVID-19 include asthma, cardiovascular disease, chronic lung disease, diabetes, chronic kidney disease, cancer in the past year, immunosuppressant therapy, and hepatitis B.

Table 2. Self-reported engagement in preventative health behaviors (N=8392).

Behavior	Value
Physical activity prevalence (N=7325), n (%)	
Physical activity performance prior to the COVID-19 pandemic	6715 (91.67)
Physical activity performance since the COVID-19 pandemic	6468 (88.30)
Physical activity change since the COVID-19 pandemic (n=6015), n (%)	
Increased	1394 (23.18)
Maintained	1585 (26.35)
Decreased	2537 (42.18)
Stopped	499 (8.30)
Meditation prevalence (n=7332), n (%)	
Meditated prior to the COVID-19 pandemic	5940 (81.01)
Meditated since the COVID-19 pandemic	5435 (74.13)
Medication change since the COVID-19 pandemic (n=5914), n (%)	
Increased	2101 (35.53)
Maintained	1979 (33.46)
Decreased	1479 (25.01)
Stopped	355 (6.00)
Strength of meditation habit (SRHI[a] score), mean (SD)	10.72 (4.05)

[a]SRHI: Self-Report Habit Index.

Table 3. Differences in mental health by region and state-level COVID-19 prevalence.

Characteristic	Stress[a]		PTSD[b,c]		Depression[d]		Anxiety[e]	
	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value	Mean (SD)	*P* value
Region[f]								
New England	18.31 (6.47)	.37	1.68 (0.89)	.82	8.92 (4.15)	.70	6.01 (3.75)	.36
Mid-Atlantic	18.66 (6.46)	<.001	1.72 (0.91)	.62	9.13 (4.22)	.03	6.07 (3.81)	.07
East North Central	17.75 (6.33)	.12	1.66 (0.90)	.56	8.81 (4.20)	.71	5.77 (3.62)	.40
West North Central	17.92 (6.29)	.65	1.65 (0.87)	.56	8.86 (4.09)	.97	5.97 (3.82)	.51
South Atlantic	17.81 (6.31)	.14	1.63 (0.87)	.06	8.51 (4.04)	<.001	5.73 (3.57)	.18
East South Central	18.76 (6.87)	.14	1.67 (0.93)	.92	9.15 (4.43)	.34	5.92 (3.70)	.81
West South Central	17.73 (6.43)	.24	1.62 (0.92)	.14	8.86 (4.24)	.96	5.67 (3.63)	.22
Mountain West	18.03 (6.13)	.94	1.69 (0.86)	.56	8.94 (4.10)	.54	5.82 (3.47)	.73
Pacific West	18.03 (6.14)	.88	1.70 (0.87)	.16	8.88 (4.06)	.79	5.88 (3.50)	.79
State COVID-19 prevalence[g]								
High-prevalence state	18.23 (6.21)	.08	1.71 (0.86)	.004	8.89 (4.06)	.80	5.92 (3.60)	.20
Low-prevalence state	17.97 (6.38)	N/A[h]	1.65 (0.89)	N/A	8.86 (4.18)	N/A	5.81 (3.64)	N/A
Total, median (SD)	18.08 (6.31)	N/A	1.67 (0.89)	N/A	8.72 (4.13)	N/A	5.86 (3.62)	N/A

[a]Stress was measured using the total score on the Perceived Stress Scale.

[b]PTSD: posttraumatic stress disorder.

[c]PTSD was measured using mean scores on the Impact of Events Scale-6.

[d]Depression was measured using the Depression subscale score on the Hospital Depression and Anxiety Scale.

[e]Anxiety was measured using the Anxiety subscale score on the Hospital Depression and Anxiety Scale.

[f]States within each region are based on US Census divisions.

[g]State-level COVID-19 prevalence is based on the number of COVID-19 cases per 100,000 at the time of the first survey distribution (April 22, 2020); high-prevalence states were California, Colorado, Illinois, Massachusetts, New Jersey, New York, and Washington.

[h]N/A: not applicable.

Table 4. Differences in COVID-19–related worry, attention to news, and stress by region and state-level COVID-19 prevalence.

Variable	Worry		High attention to news		High stress from social distancing	
	Mean (SD)	*P* value	n (%)	*P* value	n (%)	*P* value
Region						
New England	9.69 (2.85)	.54	155 (31.70)	.36	233 (47.65)	.12
Mid-Atlantic	10.11 (2.86)	<.001	353 (37.35)	.01	468 (49.47)	<.001
East North Central	9.58 (2.94)	.70	312 (32.74)	.53	404 (42.44)	.26
West North Central	9.55 (2.78)	.61	153 (33.33)	.91	213 (46.41)	.35
South Atlantic	9.59 (2.92)	.74	395 (33.14)	.69	494 (41.44)	.39
East South Central	9.75 (3.01)	.53	69 (34.85)	.78	82 (41.41)	.47
West South Central	9.32 (3.08)	.03	162 (31.03)	.20	209 (40.04)	.053
Mountain West	9.41 (2.99)	.052	225 (31.56)	.22	300 (42.08)	.25
Pacific West	9.52 (2.90)	.16	544 (34.85)	.28	706 (45.11)	.42
State COVID-19 prevalence[a]						
High-prevalence state	9.72 (2.90)	.01	1064 (34.81)	.04	1420 (46.37)	.004
Low-prevalence state	9.53 (2.95)	N/A[b]	1387 (32.51)	N/A	1832 (42.92)	N/A
Total	9.61 (2.93)	N/A	2451 (33.47)	N/A	3252 (44.37)	N/A

[a]State-level COVID-19 prevalence is based on the number of COVID-19 cases per 100,000 at the time of the baseline survey (April 22, 2020); high-prevalence states were California, Colorado, Illinois, Massachusetts, New Jersey, New York, and Washington.

[b]N/A: not applicable.

Associations Between Health Behavior Engagement and COVID-19–Related Worry, Attention to News, and Stress

Participants who were more worried about COVID-19, paid more attention to COVID-19 news and updates, and experienced more stress due to COVID-19 social distancing recommendations had greater decreases in physical activity and lower strength of their meditation habits (see Table 5). Attention to news and updates about COVID-19 and stress due to social distancing recommendations were also associated with stopping meditation.

Men, White and non-Hispanic respondents, and respondents with higher levels of education and higher household incomes were also less likely to decrease their engagement in physical activity. Strength of meditation habit was also generally greater among respondents who were older, White, non-Hispanic, female, and more educated, and who had higher annual household incomes. Younger participants and men were more likely to report that since the COVID-19 pandemic, they had stopped meditating. Living in a state with high COVID-19 prevalence was associated with decreases in physical activity and lower strength of meditation habit but not with stopping meditation practice.

Table 5. Associations between health behavior engagement and COVID-19–related worry, attention to news, and stress. The table presents the coefficients from ordinary least squares regression models for the continuous outcomes, changes in physical activity, and meditation habit strength on COVID-19–related worry, attention to news, and stress, and logistic regression models for the stopped meditation outcome on COVID-19–related worry, attention to news, and stress. Standard errors were estimated using heteroscedasticity-robust procedures.

Outcome and covariates	COVID-19 worry		COVID-19 attention		Stress about social distancing	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
Outcome: changes in physical activity						
COVID-19 worry	–0.04 (0.01)	<.001	—[a]	—	—	—
COVID-19 attention	—	—	–0.16 (0.06)	.01	—	—
Stress about social distancing	—	—	—	—	–0.16 (0.05)	.003
Demographic covariates[b]						
Age	–0.01 (0.002)	.01	–0.003 (0.002)	.09	–0.005 (0.002)	.02
Racial minority status	0.02 (0.10)	.81	0.05 (0.10)	.60	0.05 (0.10)	.60
Female	0.25 (0.07)	<.001	0.22 (0.07)	.002	0.24 (0.07)	.001
Hispanic	–0.20 (0.11)	.07	–0.26 (0.11)	.02	–0.25 (0.11)	.03
High school education only	–0.71 (0.19)	<.001	–0.70 (0.19)	<.001	–0.69 (0.19)	<.001
Undergraduate education	–0.22 (0.06)	<.001	–0.22 (0.06)	<.001	–0.22 (0.06)	<.001
Income <US $80,000	–0.29 (0.07)	<.001	–0.29 (0.07)	<.001	–0.28 (0.07)	<.001
Income of US $81,000-100,000	–0.16 (0.06)	.01	–0.17 (0.06)	.01	–0.16 (0.06)	.01
Unemployed	–0.08 (0.09)	.40	–0.07 (0.09)	.42	–0.07 (0.09)	.46
Underlying medical condition	–0.12 (0.06)	.03	–0.15 (0.06)	.01	–0.16 (0.06)	.01
Living in a state with high COVID-19 prevalence[c]	–0.01 (0.05)	.04	–0.11 (0.054)	.03	–0.12 (0.05)	.03
Outcome: stopped meditation						
COVID-19 worry	0.04 (0.02)	.045	—	—	—	—
COVID-19 attention	—	—	0.36 (0.13)	.004	—	—
Stress about social distancing	—	—	—	—	0.31 (0.12)	.01
Outcome: strength of meditation habit						
COVID-19 worry	–0.06 (0.02)	.001	—	—	—	—
COVID-19 attention	—	—	–0.17 (0.12)	.18	—	—
Stress about social distancing	—	—	—	—	–0.45 (0.11)	<.001

[a]—: not applicable.

[b]These covariates were included in all the models.

[c]State-level COVID-19 prevalence is based on the number of COVID-19 cases per 100,000 at the time of the first survey distribution (April 22, 2020); high-prevalence states were California, Colorado, Illinois, Massachusetts, New Jersey, New York, and Washington.

Mediation Models

Perceived Stress and COVID-19–Related Worry, Attention to News, and Stress via Engagement in Health Behaviors

The mediating effect of health behavior changes on the associations between stress and COVID-19–related worry, attention to news, and stress from social distancing is demonstrated by the regression analyses presented in Table 6. The first row and first column of Table 6 present the total effect of worry about COVID-19 on self-reported stress, and the subsequent columns illustrate how the association between worry about COVID-19 and stress is attenuated by the inclusion of each behavioral change measure. Based on bootstrapped estimation procedures, we found that the association between worry about COVID-19 and perceived stress was significantly mediated by changes in physical activity (*P*<.001), stopping meditation (*P*=.046), and strength of meditation habit (*P*<.001). The second panel of Table 6 presents a similar mediation analysis for the association between attention to COVID-19 news and updates and perceived stress. The bootstrapped standard error calculations found that changes in physical activity (*P*=.02) and stopping meditation (*P*=.002) partially mediated the association between attention to COVID-19 news and updates and perceived stress; however, strength of meditation habit was not a significant mediator (*P*=.11). Finally, the association between perceived stress and stress due to

COVID-19 social distancing recommendations was significantly mediated by changes in physical activity (*P*=.01), stopping meditation (*P*=.02), and strength of meditation habit (*P*<.001).

Table 6. Mediating effect of health behavior change on stress. The table presents the coefficients from ordinary least squares regression models of PSS stress score on COVID-19 worry, COVID-19 attention, and stress from social distancing as well as the indicated health behavior changes. The demographic variables of age, male sex, Hispanic, income <US $80,000, income US $81,000-100,000, unemployed, and underlying medical condition were included as covariates in all models, which also estimated heteroscedasticity-robust standard errors.

Variable	PSS[a] stress score							
	Total effect		Change in physical activity		Stopped meditating		Strength of meditation habit	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
COVID-19 worry—total effect	0.61 (0.02)	<.001	—[b]	—	—	—	—	—
COVID-19 worry—indirect effect			0.60 (0.03)	<.001	0.64 (0.028)	<.001	0.63 (0.03)	<.001
Change in physical activity	—	—	–0.25 (0.03)	<.001	—	—	—	—
Stopped meditating	—	—	—	—	2.00 (0.33)	<.001	—	—
Strength of meditation habit	—	—	—	—	—	—	–0.15 (0.02)	<.001
COVID-19 attention—total effect	0.77 (0.16)	<.001	—	—	—	—	—	—
COVID-19 attention—indirect effect			0.74 (0.16)	<.001	0.65 (0.18)	<.001	0.76 (0.18)	<.001
Change in physical activity	—	—	–0.30 (0.04)	<.001	—	—	—	—
Stopped meditating	—	—	—	—	2.15 (0.35)	<.001	—	—
Strength of meditation habit	—	—	—	—	—	—	–0.17 (0.02)	<.001
Stress from social distancing—total effect	3.59 (0.16)	<.001	—	—	—	—	—	—
Stress from social distancing—indirect effect			3.54 (0.14)	<.001	3.64 (0.16)	<.001	3.59 (0.16)	<.001
Change in physical activity	—	—	–0.27 (0.03)	<.001	—	—	—	—
Stopped meditating	—	—	—	—	1.93 (0.33)	<.001	—	—
Strength of meditation habit	—	—	—	—	—	—	–0.15 (0.02)	<.001

[a]PSS: Perceived Stress Scale.

[b]—: not applicable.

PTSD Symptoms and COVID-19–Related Worry, Attention to News, and Stress via Engagement in Health Behaviors

The mediating effect of health behavior changes on the associations between PTSD symptoms and COVID-19–related worry, attention to news, and stress is outlined by the regression analyses presented in Table 7. The first row and first column of Table 7 present the total effect of worry about COVID-19 on self-reported PTSD symptoms, and the subsequent columns illustrate how the association between worry about COVID-19 and PTSD symptoms is attenuated by the inclusion of each behavioral change measure. Based on bootstrapped estimation procedures, we found that the association between worry about COVID-19 and PTSD symptoms was significantly mediated by changes in physical activity (*P*<.001) and strength of meditation habit (*P*<.001) but not significantly mediated by stopping meditation (*P*=.07). The second panel of Table 7 presents the same mediation analysis for the association between attention to COVID-19 news and updates and PTSD symptoms, where bootstrapped standard error calculations indicated that changes in physical activity (*P*=.01) and stopping meditation (*P*=.004) significantly mediated the association between PTSD symptoms and attention to COVID-19 news and updates while the strength of meditation habit did not (*P*=.13). Finally, the association between PTSD symptoms and stress caused by COVID-19 social distancing recommendations was partially mediated by changes in physical activity (*P*=.01), stopping meditation (*P*=.01), and strength of meditation habit (*P*<.001).

Table 7. Mediating effect of health behavior change on PTSD symptoms. This table presents the coefficients from ordinary least squares regression models of PTSD score on COVID-19–related worry, attention to news, and stress from social distancing as well as the indicated health behavior changes. Age, male sex, Hispanic race, income <US $80,000, income US $81,000-100,000, unemployed, and underlying medical condition were included as covariates in all models.

Variable	PTSD[a,b] score							
	Total effect		Change in physical activity		Stopped meditating		Strength of meditation habit	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
COVID-19 worry—total effect	0.13 (0.003)	<.001	—[c]	—	—	—	—	—
COVID-19 worry—indirect effect			0.13 (0.003)	<.001	0.13 (0.00)	<.001	0.13 (0.004)	<.001
Change in physical activity	—	—	–0.03 (0.010)	<.001	—	—	—	—
Stopped meditating	—	—	—	—	0.23 (0.05)	<.001		—
Strength of meditation habit	—	—	—	—	—	—	–0.01 (0.003)	<.001
COVID-19 attention—total effect	0.26 (0.02)	<.001	—	—	—	—	—	—
COVID-19 attention—indirect effect			0.26 (0.02)	<.001	0.24 (0.03)	<.001	0.25 (0.03)	<.001
Change in physical activity	—	—	–0.03 (0.010)	<.001	—	—	—	—
Stopped meditating	—	—	—	—	0.25 (0.05)	<.001	—	—
Strength of meditation habit	—	—	—	—	—	—	–0.02 (0.003)	<.001
Stress from social distancing—total effect	0.55 (0.02)	<.001	—	—	—	—	—	—
Stress from social distancing-indirect effect			0.55 (0.02)	<.001	0.57 (0.02)	<.001	0.56 (0.02)	<.001
Change in physical activity	—	—	–0.03 (0.010)	<.001	—	—	—	—
Stopped meditating	—	—	—	—	0.23 (0.05)	<.001	—	—
Strength of meditation habit	—	—	—	—	—	—	–0.01 (0.003)	<.001

[a]PTSD: posttraumatic stress disorder.

[b]PTSD was measured using mean scores on the Impact of Events Scale-6.

[c]—: not applicable.

Discussion

Our findings describe the mental health impact of the COVID-19 pandemic and outline how continued participation in health behaviors such as physical activity and mindfulness meditation reduce worsening of mental health due to the COVID-19 pandemic. The aim of this baseline survey was to first examine the regional differences in mental health and COVID-19–related worry, attention to news, and stress in light of the state-level prevalence of COVID-19 infections at the time of this survey. We additionally sought to estimate the associations between COVID-19–related worry and attention to news and stress, health behavior engagement, and mental health, as well as to explore the mediating effect of health behavior engagement on the associations between mental health and COVID-19–related worry, attention to news, and stress.

Mental Health and COVID-19–Related Worry, Attention to News, and Stress by Region

Our findings indicate that participants living in the Mid-Atlantic region (ie, New Jersey, New York, or Pennsylvania) had higher levels of stress and more severe depressive symptoms compared to the rest of the country, which is consistent with the fact that a high prevalence of confirmed COVID-19 cases existed in this region during the time period when this survey was administered. Participants in the Mid-Atlantic region were also more worried about COVID-19, paid more attention to COVID-19 news and updates, and reported more stress related to social distancing recommendations than survey participants living in other regions. Additionally, participants living in states with a high prevalence of COVID-19 cases (ie, California, Colorado, Illinois, Massachusetts, New Jersey, New York, or Washington) reported more severe PTSD symptoms than those living in states where COVID-19 was less prevalent, and they also reported more COVID-19–related worry, paid more

attention to COVID-19 news and updates, and experienced more stress from social distancing recommendations. These preliminary findings highlight the regional differences in worry and fear about COVID-19 and mental health based on the prevalence of COVID-19 cases. As the prevalence of COVID-19 cases among regions may change over time, this information could be particularly useful for public health agencies or community health centers to direct or provide more mental health resources to regions with high COVID-19 prevalence in an effort to mitigate the subsequent mental health burden. Additionally, future research and public health interventions should consider the local prevalence of COVID-19 and the extent to which individuals are paying attention to news and media to monitor news related to COVID-19 when targeting mental health interventions, as these factors may play an important role in the mental health impact of the pandemic.

Associations of Worry About COVID-19, Physical Activity, Meditation, and Mental Health

Our findings suggest that higher levels of worry about COVID-19 were associated with lower physical activity levels and lower strength of meditation habits; however, overall, the majority of participants reported increasing or maintaining their physical activity or meditation practice. Although our findings are not causal, higher levels of worry about COVID-19 may have contributed to the difficulty people had in sustaining their physical activity participation and engaging in habitual meditation practice. This is aligned with other research demonstrating that anxiety and worry can be detrimental to habitual participation in health-promoting behaviors, both in general [48] and in the context of COVID-19 [49,50]. One cross-sectional survey conducted in Belgium during the COVID-19 lockdown reported increases in physical activity, but only in those who were previously less active; decreases in physical activity were found in previously highly active adults [51]. Conversely, other cross-sectional studies conducted during the COVID-19 pandemic reported significant declines in physical activity and increased sitting time [49,52,53]. No studies have reported patterns of meditation practice; however, one cross-sectional study assessing stress-coping behaviors conducted among New York City–based health care workers during the COVID-19 pandemic reported that meditation (23%) was a commonly endorsed behavior, along with physical activity/exercise (59%) [54]. However, the physical activity and meditation patterns were not assessed prior to the COVID-19 pandemic. More research describing patterns of health behaviors during the COVID-19 pandemic is needed to better understand how COVID-19 may contribute to long-lasting, negative health behavior changes.

Our data also suggest that higher levels of worry about COVID-19 are associated with poor mental health (ie, increased stress and PTSD), and importantly, this association between worry about COVID-19 and poor mental health was mediated by lower physical activity levels and lower strength of meditation habits (similar associations were observed for depression and anxiety; see Multimedia Appendix 1). Therefore, stopping or reducing health promoting behaviors, particularly physical activity and meditation practice, during the COVID-19 pandemic may increase the negative impact of COVID-19–related worry on stress, PTSD, depression, and anxiety. To our knowledge, this is the first study to examine the mediating effects of physical activity and meditation on the associations between worry about COVID-19 and mental health during the pandemic. Existing research has shown that greater concern about COVID-19 was associated with greater anxiety and depression levels [55], and decreased physical activity and increased sedentary time during COVID-19 was associated with poor mental health (ie, higher stress, anxiety, and depression symptoms) [49,56]. Additionally, both physical activity and meditation have been recommended as healthy ways to cope with stress during the COVID-19 pandemic, underscoring the importance of maintaining these behaviors during a time of heightened stress [57]. Our data demonstrate that participation in physical activity and meditation are important mechanisms in reducing the increased mental health problems associated with COVID-19–related worry. Other studies have shown that both mindfulness-based and physical activity interventions may be protective against the development of trauma-related psychopathology (eg, PTSD) by enhancing cognitive function, reducing arousal, normalizing hypothalamic pituitary axis function, and reducing inflammatory markers [58,59]. There is a need for more research, especially longitudinal data, to examine the long-term changes in health behaviors due to the COVID-19 pandemic, better disentangle the causal relationships between these psychological and behavioral outcomes, and identify strategies for helping people maintain health-promoting behaviors.

In summary, our findings support existing evidence of the beneficial health effects of physical activity and meditation on mental health outcomes [16,60-62]. Because the COVID-19 pandemic has profoundly impacted daily routines (eg, social distancing, quarantine, businesses closures) and may negatively impact the performance of health behaviors, it is important to continue promoting self-management of health behaviors such as physical activity and meditation that can reduce worsening of mental health during the COVID-19 pandemic, particularly in regions with a heightened sense of worry about COVID-19. Encouraging participation in physical activity and meditation should be an important public health objective during the current COVID-19 pandemic, especially because reduced physical health and poor mental health have been shown to increase susceptibility to COVID-19 infection and disease transmission [63,64]. Public health agencies may consider providing strategies to help people maintain or adapt their current health behaviors. For example, digital or mHealth interventions for both physical activity and meditation have shown promise for their feasibility, scalability, and physical and mental health benefits [64-67]. Digital and mHealth interventions are also convenient and often budget-friendly ways to encourage participation in both physical activity and meditation, and more research is needed to better understand their efficacy and applicability during the COVID-19 pandemic.

Limitations

Although this study is one of the first to describe the associations between COVID-19–related worry and attention to news and stress, mental health, and self-management health behaviors, there are important limitations to be noted. First, our sample

was primarily female, non-Hispanic, White, high-income, and highly educated, and the participants were paid subscribers of Calm, which limits the generalizability of these data. Second, this survey was cross-sectional; therefore, causal relationships cannot be determined from these analyses. In the broader study, we plan to implement four more surveys over the next 12 months to provide a more comprehensive longitudinal assessment of the impact of the COVID-19 pandemic on mental health and health behaviors.

Conclusions

Our findings underscore the importance of maintaining self-management health behaviors such as physical activity and meditation for sustaining one's mental health during the COVID-19 pandemic. These results suggest that public health agencies and health organizations should promote the maintenance of health habits with strategies such as digital and mHealth approaches that can be more easily adapted during stay-at-home orders and social distancing mandates. Future research is needed to identify the causal relationships between these psychological and behavioral outcomes and to evaluate strategies for helping people adapt their current meditation and physical activity practices to the restrictions on daily life imposed by COVID-19–related public health policies.

Conflicts of Interest

JH is currently the Director of Science at Calm. JH had been conducting research with Calm as a partner for almost 5 years before becoming the Director of Science and the Scientific Advisory Board. Her role is to ensure the quality of Calm's science. There were no financial incentives from the growth of Calm for any author.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 21 KB - mental_v8i4e28479_app1.docx]

References

1. Mental health and psychosocial considerations during the COVID-19 outbreak. World Health Organization. 2020 Mar 18. URL: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf [accessed 2021-04-05]
2. Fiorillo A, Gorwood P. The consequences of the COVID-19 pandemic on mental health and implications for clinical practice. Eur Psychiatry 2020 Apr 01;63(1):e32 [FREE Full text] [doi: 10.1192/j.eurpsy.2020.35] [Medline: 32234102]
3. Bao Y, Sun Y, Meng S, Shi J, Lu L. 2019-nCoV epidemic: address mental health care to empower society. Lancet 2020 Feb 22;395(10224):e37-e38 [FREE Full text] [doi: 10.1016/S0140-6736(20)30309-3] [Medline: 32043982]
4. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. Asian J Psychiatr 2020 Aug;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
5. Fitzpatrick KM, Harris C, Drawve G. Fear of COVID-19 and the mental health consequences in America. Psychol Trauma 2020 Aug;12(S1):S17-S21. [doi: 10.1037/tra0000924] [Medline: 32496100]
6. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
7. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
8. Salisbury H. Helen Salisbury: Fear in the time of covid. BMJ 2020 Mar 30;368:m1286. [doi: 10.1136/bmj.m1286] [Medline: 32229528]
9. Thombs B, Tao L, Wu Y, Levis B, Sun Y, Bourgeault A, et al. Preliminary COVID-19 Fears Questionnaire: systemic sclerosis and chronic medical conditions versions. OSF Preprints. Preprint posted on April 6, 2020. [doi: 10.31219/osf.io/m2ybt]
10. Thombs BD, Bonardi O, Rice DB, Boruff JT, Azar M, He C, et al. Curating evidence on mental health during COVID-19: a living systematic review. J Psychosom Res 2020 Jun;133:110113 [FREE Full text] [doi: 10.1016/j.jpsychores.2020.110113] [Medline: 32354463]
11. 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] [Medline: 32304649]
12. Inkster B, O'Brien R, Selby E, Joshi S, Subramanian V, Kadaba M, et al. Digital health management during and beyond the COVID-19 pandemic: opportunities, barriers, and recommendations. JMIR Ment Health 2020 Jul 06;7(7):e19246 [FREE Full text] [doi: 10.2196/19246] [Medline: 32484783]
13. Lean M, Fornells-Ambrojo M, Milton A, Lloyd-Evans B, Harrison-Stewart B, Yesufu-Udechuku A, et al. Self-management interventions for people with severe mental illness: systematic review and meta-analysis. Br J Psychiatry 2019 May 22;214(5):260-268 [FREE Full text] [doi: 10.1192/bjp.2019.54] [Medline: 30898177]

14. Edenfield TM, Saeed SA. An update on mindfulness meditation as a self-help treatment for anxiety and depression. PRBM 2012 Oct:131. [doi: 10.2147/prbm.s34937]
15. Keng S, Smoski MJ, Robins CJ. Effects of mindfulness on psychological health: a review of empirical studies. Clin Psychol Rev 2011 Aug;31(6):1041-1056 [FREE Full text] [doi: 10.1016/j.cpr.2011.04.006] [Medline: 21802619]
16. Paluska SA, Schwenk TL. Physical activity and mental health: current concepts. Sports Med 2000 Mar;29(3):167-180. [doi: 10.2165/00007256-200029030-00003] [Medline: 10739267]
17. Beaulac J, Carlson A, Boyd RJ. Counseling on physical activity to promote mental health: practical guidelines for family physicians. Can Fam Physician 2011 Apr;57(4):399-401 [FREE Full text] [Medline: 21490348]
18. Saxena S, Van Ommeren M, Tang KC, Armstrong TP. Mental health benefits of physical activity. J Ment Health 2009 Jul 06;14(5):445-451. [doi: 10.1080/09638230500270776]
19. Physical activity and health: a report of the Surgeon General. US Centers for Disease Control and Prevention. 1999. URL: https://www.cdc.gov/nccdphp/sgr/index.htm [accessed 2021-04-06]
20. Hall G, Laddu DR, Phillips SA, Lavie CJ, Arena R. A tale of two pandemics: how will COVID-19 and global trends in physical inactivity and sedentary behavior affect one another? Prog Cardiovasc Dis 2021;64:108-110 [FREE Full text] [doi: 10.1016/j.pcad.2020.04.005] [Medline: 32277997]
21. Burtscher J, Burtscher M, Millet GP. (Indoor) isolation, stress, and physical inactivity: vicious circles accelerated by COVID-19? Scand J Med Sci Sports 2020 Aug 06;30(8):1544-1545 [FREE Full text] [doi: 10.1111/sms.13706] [Medline: 32374894]
22. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. J Behav Med 2008 Feb;31(1):23-33. [doi: 10.1007/s10865-007-9130-7] [Medline: 17899351]
23. Lacaille J, Sadikaj G, Nishioka M, Carrière K, Flanders J, Knäuper B. Daily Mindful Responding Mediates the Effect of Meditation Practice on Stress and Mood: The Role of Practice Duration and Adherence. J Clin Psychol 2018 Jan 31;74(1):109-122. [doi: 10.1002/jclp.22489] [Medline: 28759111]
24. Wood W. Habit in Personality and Social Psychology. Pers Soc Psychol Rev 2017 Nov 24;21(4):389-403. [doi: 10.1177/1088868317720362] [Medline: 28737111]
25. Wood W, Rünger D. Psychology of habit. Annu Rev Psychol 2016;67:289-314. [doi: 10.1146/annurev-psych-122414-033417] [Medline: 26361052]
26. Verplanken B, Orbell S. Reflections on past behavior: a self‐report index of habit strength. J Appl Soc Psychol 2003 Jun;33(6):1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
27. Behan C. The benefits of meditation and mindfulness practices during times of crisis such as COVID-19. Ir J Psychol Med 2020 Dec 14;37(4):256-258 [FREE Full text] [doi: 10.1017/ipm.2020.38] [Medline: 32406348]
28. 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 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
29. Zhou X, Snoswell CL, Harding LE, Bambling M, Edirippulige S, Bai X, et al. The role of telehealth in reducing the mental health burden from COVID-19. Telemed J E Health 2020 Apr 23;26(4):377-379 [FREE Full text] [doi: 10.1089/tmj.2020.0068] [Medline: 32202977]
30. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. JMIR mHealth uHealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
31. Yang E, Schamber E, Meyer RML, Gold JI. Happier healers: randomized controlled trial of mobile mindfulness for stress management. J Altern Complement Med 2018 May;24(5):505-513. [doi: 10.1089/acm.2015.0301] [Medline: 29420050]
32. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: a randomized controlled trial. Mindfulness (N Y) 2018 Mar 1;9(5):1584-1593 [FREE Full text] [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
33. Puzia ME, Huberty J, Eckert R, Larkey L, Mesa R. Associations between global mental health and response to an app-based meditation intervention in myeloproliferative neoplasm patients. Integr Cancer Ther 2020 Jun 20;19:1534735420927780 [FREE Full text] [doi: 10.1177/1534735420927780] [Medline: 32564631]
34. Huberty J, Eckert R, Larkey L, Kurka J, Rodríguez De Jesús SA, Yoo W, et al. Smartphone-based meditation for myeloproliferative neoplasm patients: feasibility study to inform future trials. JMIR Form Res 2019 Apr 29;3(2):e12662 [FREE Full text] [doi: 10.2196/12662] [Medline: 31033443]
35. Calm COVID-19 survey. OSF. URL: https://osf.io/dsgrv/ [accessed 2021-04-06]
36. Form: CDC COVID-19 community survey question bank (DRAFT). US National Library of Medicine. URL: https://cde.nlm.nih.gov/formView?tinyId=Kcceysolt [accessed 2021-04-06]
37. Vartti A, Oenema A, Schreck M, Uutela A, de Zwart O, Brug J, et al. SARS knowledge, perceptions, and behaviors: a comparison between Finns and the Dutch during the SARS outbreak in 2003. Int J Behav Med 2009 Jan 31;16(1):41-48 [FREE Full text] [doi: 10.1007/s12529-008-9004-6] [Medline: 19184625]

38. ZIP code data. US General Services Administration. 2020. URL: https://catalog.data.gov/dataset/zip-code-data [accessed 2020-01-06]
39. Census Regions and Divisions of the United States. US Census Bureau. URL: https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf [accessed 2021-04-06]
40. csse_covid_19_daily_reports_us. GitHub. URL: https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_daily_reports_us [accessed 2021-04-06]
41. Gardner B, Abraham C, Lally P, de Bruijn G. Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. Int J Behav Nutr Phys Act 2012 Aug 30;9:102 [FREE Full text] [doi: 10.1186/1479-5868-9-102] [Medline: 22935297]
42. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385. [doi: 10.2307/2136404]
43. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Claremont Symposium on Applied Social Psychology. The Social Psychology of Health. Thousand Oaks, CA: Sage Publications; 1988:31-67.
44. Lee E. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci) 2012 Dec;6(4):121-127 [FREE Full text] [doi: 10.1016/j.anr.2012.08.004] [Medline: 25031113]
45. Hosey MM, Leoutsakos JS, Li X, Dinglas VD, Bienvenu OJ, Parker AM, et al. Screening for posttraumatic stress disorder in ARDS survivors: validation of the Impact of Event Scale-6 (IES-6). Crit Care 2019 Aug 07;23(1):276 [FREE Full text] [doi: 10.1186/s13054-019-2553-z] [Medline: 31391069]
46. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
47. The R Project for statistical computing. URL: https://www.r-project.org/ [accessed 2021-04-06]
48. Banerjee M, Cavanagh K, Strauss C. Barriers to mindfulness: a path analytic model exploring the role of rumination and worry in predicting psychological and physical engagement in an online mindfulness-based intervention. Mindfulness (N Y) 2018;9(3):980-992 [FREE Full text] [doi: 10.1007/s12671-017-0837-4] [Medline: 29875884]
49. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. Int J Environ Res Public Health 2020 Jun 07;17(11):4065 [FREE Full text] [doi: 10.3390/ijerph17114065] [Medline: 32517294]
50. Schuch FB, Bulzing RA, Meyer J, Vancampfort D, Firth J, Stubbs B, et al. Associations of moderate to vigorous physical activity and sedentary behavior with depressive and anxiety symptoms in self-isolating people during the COVID-19 pandemic: a cross-sectional survey in Brazil. Psychiatry Res 2020 Oct;292:113339 [FREE Full text] [doi: 10.1016/j.psychres.2020.113339] [Medline: 32745795]
51. Constandt B, Thibaut E, De Bosscher V, Scheerder J, Ricour M, Willem A. Exercising in times of lockdown: an analysis of the impact of COVID-19 on levels and patterns of exercise among adults in Belgium. Int J Environ Res Public Health 2020 Jun 10;17(11):4144 [FREE Full text] [doi: 10.3390/ijerph17114144] [Medline: 32532013]
52. Ammar A, Brach M, Trabelsi K, Chtourou H, Boukhris O, Masmoudi L, et al. Effects of COVID-19 home confinement on eating behaviour and physical activity: results of the ECLB-COVID19 international online survey. Nutrients 2020 May 28;12(6):1583 [FREE Full text] [doi: 10.3390/nu12061583] [Medline: 32481594]
53. Almandoz JP, Xie L, Schellinger JN, Mathew MS, Gazda C, Ofori A, et al. Impact of COVID-19 stay-at-home orders on weight-related behaviours among patients with obesity. Clin Obes 2020 Oct;10(5):e12386 [FREE Full text] [doi: 10.1111/cob.12386] [Medline: 32515555]
54. Shechter A, Diaz F, Moise N, Anstey DE, Ye S, Agarwal S, et al. Psychological distress, coping behaviors, and preferences for support among New York healthcare workers during the COVID-19 pandemic. Gen Hosp Psychiatry 2020 Sep;66:1-8 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.06.007] [Medline: 32590254]
55. Nelson BW, Pettitt A, Flannery JE, Allen NB. Rapid assessment of psychological and epidemiological correlates of COVID-19 concern, financial strain, and health-related behavior change in a large online sample. PLoS One 2020;15(11):e0241990 [FREE Full text] [doi: 10.1371/journal.pone.0241990] [Medline: 33175882]
56. Cheval B, Sivaramakrishnan H, Maltagliati S, Fessler L, Forestier C, Sarrazin P, et al. Relationships between changes in self-reported physical activity, sedentary behaviour and health during the coronavirus (COVID-19) pandemic in France and Switzerland. J Sports Sci 2021 Mar;39(6):699-704. [doi: 10.1080/02640414.2020.1841396] [Medline: 33118469]
57. COVID-19: coping with stress. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html [accessed 2021-03-23]
58. Boyd JE, Lanius RA, McKinnon MC. Mindfulness-based treatments for posttraumatic stress disorder: a review of the treatment literature and neurobiological evidence. J Psychiatry Neurosci 2018 Jan 01;43(1):7-25. [doi: 10.1503/jpn.170021]
59. Hegberg NJ, Hayes JP, Hayes SM. Exercise intervention in PTSD: a narrative review and rationale for implementation. Front Psychiatry 2019 Mar 21;10:133 [FREE Full text] [doi: 10.3389/fpsyt.2019.00133] [Medline: 30949075]
60. Penedo FJ, Dahn JR. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. Curr Opin Psychiatry 2005 Mar;18(2):189-193. [doi: 10.1097/00001504-200503000-00013] [Medline: 16639173]

61. Rebar AL, Stanton R, Geard D, Short C, Duncan MJ, Vandelanotte C. A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. Health Psychol Rev 2015 Jul 03;9(3):366-378. [doi: 10.1080/17437199.2015.1022901] [Medline: 25739893]
62. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
63. Nieman DC, Wentz LM. The compelling link between physical activity and the body's defense system. J Sport Health Sci 2019 May;8(3):201-217 [FREE Full text] [doi: 10.1016/j.jshs.2018.09.009] [Medline: 31193280]
64. Morey JN, Boggero IA, Scott AB, Segerstrom SC. Current directions in stress and human immune function. Curr Opin Psychol 2015 Oct 01;5:13-17 [FREE Full text] [doi: 10.1016/j.copsyc.2015.03.007] [Medline: 26086030]
65. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR mHealth uHealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
66. Plaza I, Demarzo MMP, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. JMIR mHealth uHealth 2013 Nov 01;1(2):e24 [FREE Full text] [doi: 10.2196/mhealth.2733] [Medline: 25099314]
67. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012 Nov 21;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]

Abbreviations

CDC: US Centers for Disease Control and Prevention
HADS: Hospital Anxiety and Depression Scale
HADS-A: Hospital Anxiety and Depression Scale, anxiety subscale
HADS-D: Hospital Anxiety and Depression Scale, depression subscale
IES-6: Impact of Events Scale-6
mHealth: mobile health
PSS: Perceived Stress Scale
PTSD: posttraumatic stress disorder
SARS: severe acute respiratory syndrome
SRHI: Self-Report Habit Index

Edited by J Torous, G Eysenbach; submitted 04.03.21; peer-reviewed by M Asgari Mehrabadi, C Roos; comments to author 20.03.21; revised version received 23.03.21; accepted 24.03.21; published 13.04.21.

Please cite as:
Green J, Huberty J, Puzia M, Stecher C
The Effect of Meditation and Physical Activity on the Mental Health Impact of COVID-19–Related Stress and Attention to News Among Mobile App Users in the United States: Cross-sectional Survey
JMIR Ment Health 2021;8(4):e28479
URL: https://mental.jmir.org/2021/4/e28479
doi:10.2196/28479
PMID:33788698

Original Paper

Learning the Mental Health Impact of COVID-19 in the United States With Explainable Artificial Intelligence: Observational Study

Indra Prakash Jha[1*], MCA; Raghav Awasthi[1*], MSc; Ajit Kumar[2], MBA; Vibhor Kumar[1], PhD; Tavpritesh Sethi[1], PhD

[1]Indraprastha Institute of Information Technology, New Delhi, India

[2]Adobe, Noida, India

[*]these authors contributed equally

Corresponding Author:
Tavpritesh Sethi, PhD
Indraprastha Institute of Information Technology
Room 309, R and D Building
IIIT Campus, Okhla Phase 3
New Delhi
India
Phone: 91 01126907533
Email: tavpriteshsethi@iiitd.ac.in

Abstract

Background: The COVID-19 pandemic has affected the health, economic, and social fabric of many nations worldwide. Identification of individual-level susceptibility factors may help people in identifying and managing their emotional, psychological, and social well-being.

Objective: This study is focused on learning a ranked list of factors that could indicate a predisposition to a mental disorder during the COVID-19 pandemic.

Methods: In this study, we have used a survey of 17,764 adults in the United States from different age groups, genders, and socioeconomic statuses. Through initial statistical analysis and Bayesian network inference, we have identified key factors affecting mental health during the COVID-19 pandemic. Integrating Bayesian networks with classical machine learning approaches led to effective modeling of the level of mental health prevalence.

Results: Overall, females were more stressed than males, and people in the age group 18-29 years were more vulnerable to anxiety than other age groups. Using the Bayesian network model, we found that people with a chronic mental illness were more prone to mental disorders during the COVID-19 pandemic. The new realities of working from home; homeschooling; and lack of communication with family, friends, and neighbors induces mental pressure. Financial assistance from social security helps in reducing mental stress during the COVID-19–generated economic crises. Finally, using supervised machine learning models, we predicted the most mentally vulnerable people with ~80% accuracy.

Conclusions: Multiple factors such as social isolation, digital communication, and working and schooling from home were identified as factors of mental illness during the COVID-19 pandemic. Regular in-person communication with friends and family, a healthy social life, and social security were key factors, and taking care of people with a history of mental disease appears to be even more important during this time.

(JMIR Ment Health 2021;8(4):e25097) doi:10.2196/25097

KEYWORDS

COVID-19; mental health; Bayesian network; machine learning; artificial intelligence; disorder; susceptibility; well-being; explainable artificial intelligence

Introduction

After 7 months of initial reporting, the COVID-19 pandemic continues worldwide. The mental health consequences of the COVID-19 pandemic have been substantial. More than half a million lives and more than 400 million jobs have been lost [1], causing a considerable degree of fear, worry, and concern. These effects are seen in the population at large and may be more

pronounced among certain groups such as youth, frontline workers [2], caregivers, and people with chronic medical conditions. The new normal has introduced unprecedented interventions of countrywide lockdowns that are necessary to control the spread but have led to increased social isolation. Loneliness, depression, harmful alcohol and drug use, and self-harm or suicidal behavior are also expected to rise.

The Lancet Psychiatry [3] recently highlighted the needs of vulnerable groups during this time, including those with severe mental illness, learning difficulties, and neurodevelopmental disorders, as well as socially excluded groups such as prisoners, the homeless, and refugees. Calls to action for engaging more early-career psychiatrists [4,5], using technology such as telepsychiatry, and stressing the high susceptibility of frontline medical workers themselves [6] have highlighted the magnitude of the problem. Further, interventions are expected to have a gender-specific impact, with women more likely to be exposed to additional stressors related to informal care, already existing economic disparity, and school closures. Similarly, age and comorbidity status may have a direct impact on susceptibility to mental health challenges due to their relationship with COVID-19 morbidity and mortality. Indeed, it has been established that emotional distress is ubiquitous in affected populations—a finding certain to be echoed in populations affected by the COVID-19 pandemic [7]. Finally, the role of social media [8,9] is complex, with some research indicating an association between social media exposure and a higher prevalence of mental health problems [10].

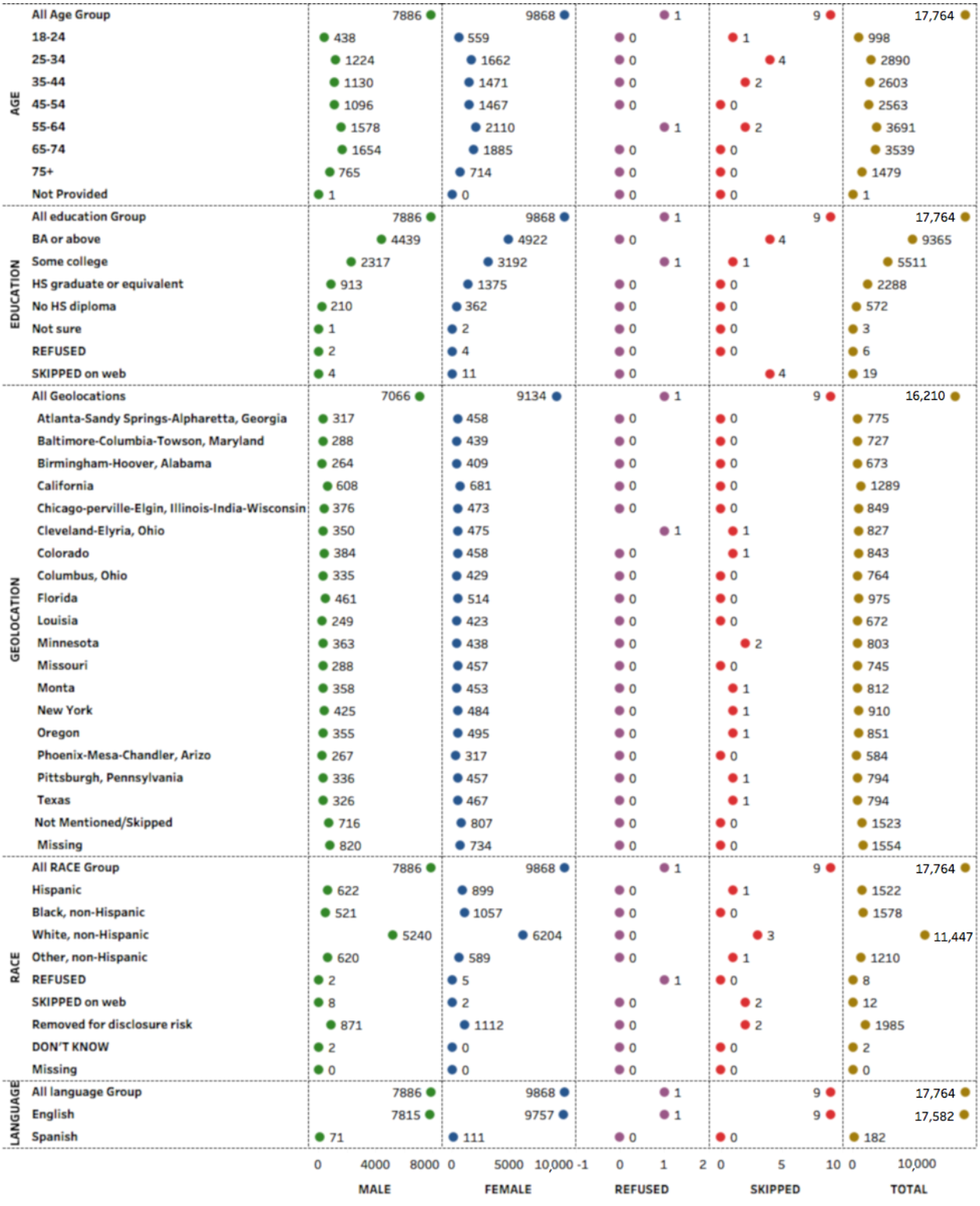
However, most of these effects have been studied in isolation with a lack of modeling the collective impact of such factors. This study addresses this gap through the use of Bayesian networks (BNs), an explainable artificial intelligence approach that captures the joint multivariate distribution underlying large survey data collected across the United States. We also address the gap of vulnerability prediction for mental health events such as anxiety attacks using supervised machine learning models.

Methods

Data Sets

We extracted the data of 17,764 adults [11] from two weekly surveys (April 20-26 and May 4-10, 2020) of the US adult household population nationwide for 18 regional areas including 10 states (California, Colorado, Florida, Louisiana, Minnesota, Missouri, Montana, New York, Oregon, Texas) and 8 metropolitan statistical areas (Atlanta, Baltimore, Birmingham, Chicago, Cleveland, Columbus, Phoenix, Pittsburgh). Two rounds of data collection were available at the time of this analysis, and both rounds of data until May 25, 2020, were included in this analysis. The details of the original data are available elsewhere [12]. To summarize, the data set comprised variables on physical health, mental health, insurance-related policy, economic security, and social dynamics. Figure 1 shows the sociodemographic characteristics of respondents participating in the survey.

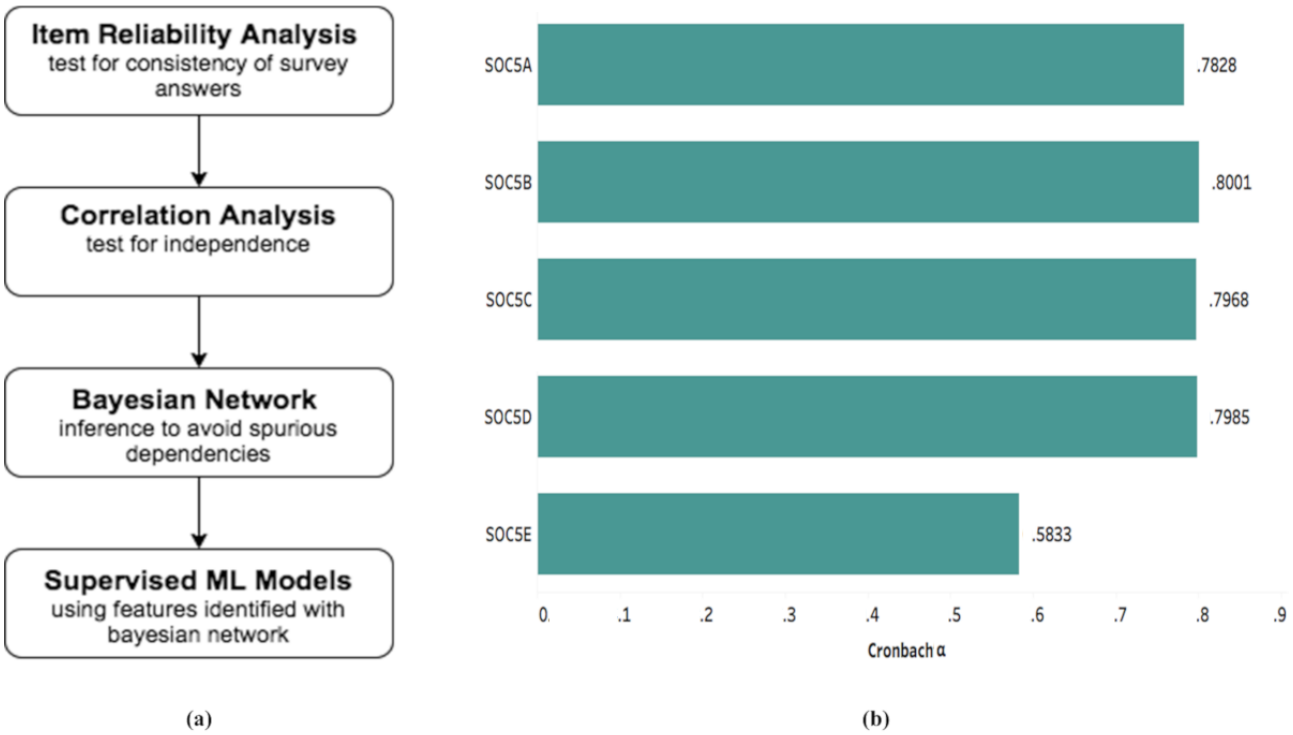
Figure 1. Sociodemographics of respondents who participated in the survey. It can be seen that there was almost a similar representation from both genders. Age groups 25-75 years were predominantly captured in the survey. Most of the respondents had received a Bachelor's degree or above and were nearly equally distributed across geographies within the United States. HS: high school.



Analysis

Figure 2a shows the flow diagram for the analyses conducted. The survey questions were classified into several types of indicators such as mental health, work from home, communication, COVID-19 symptoms, chronic medical conditions, behavioral aspects, insurance assistance, and many others (Multimedia Appendix 1).

Figure 2. (a) Outline of the analytical pipeline. (b) Item reliability analysis of mental health indicators revealed a high degree of internal consistency (Cronbach alpha value >.70) for most of the psychological variables, thus indicating suitability for the modeling exercise. ML: machine learning.



Item Reliability Analysis

We constructed a model for the mental health indicators with attribute *soc5a* (felt nervous, anxious, or on edge), attribute *soc5b* (felt depressed), attribute *soc5c* (felt lonely), attribute *soc5d* (felt hopeless about the future), and *soc5e* (sweating, trouble breathing, pounding heart, etc in the last 7 days) as outcome variables. Hence, we first evaluated the consistency in answers to the mental health questions using an item reliability analysis. A scale for measuring the reliability of internal consistency, Cronbach alpha, was calculated using the *Psych* package in R (R Foundation for Statistical Computing) [13].

Test of Independence Among the Mental Health Indicator and Other Indicators

Thereafter, a pairwise chi-square test of independence was performed to examine associations between *mental health indicators* and other variables, and a *P* value<.05 was taken as the cutoff for significance.

Data-Driven Bayesian Network Analysis

Since mental health variables may have complex dependencies with potential confounding factors, mediation, and intercausal dependency, we extended our association analysis with data-driven BN structure learning. The structure of the learned BN was made robust through bootstrapping and ensemble averaging of edge directions. The hill climbing optimizer [14] with the Akaike information criterion–based score [15] was used to select the best probabilistic graphical model that explained the data. Bootstrapped learning and majority voting over 101 BNs were done. Exact inference using the belief propagation algorithm [16] was learned to quantify the strength of learned associations. The analysis was performed in R using the package *wiseR* [17].

Mental Health Prediction Using Supervised Machine Learning

Next, the *Markov blanket* [18] of *mental health indicators* was extracted to select features that may predict responses to the *mental health indicators*. Data were partitioned into training (80%) and testing (20%) sets, and the class imbalance was corrected using the synthetic minority oversampling technique [19]. Different supervised machine learning models—random forest (RF), support vector machine (SVM), logistic regression, naive Bayes—were learned for predicting the response to mental health indicators using the Scikit-learn library [20] in Python.

Results

Item Reliability Analysis

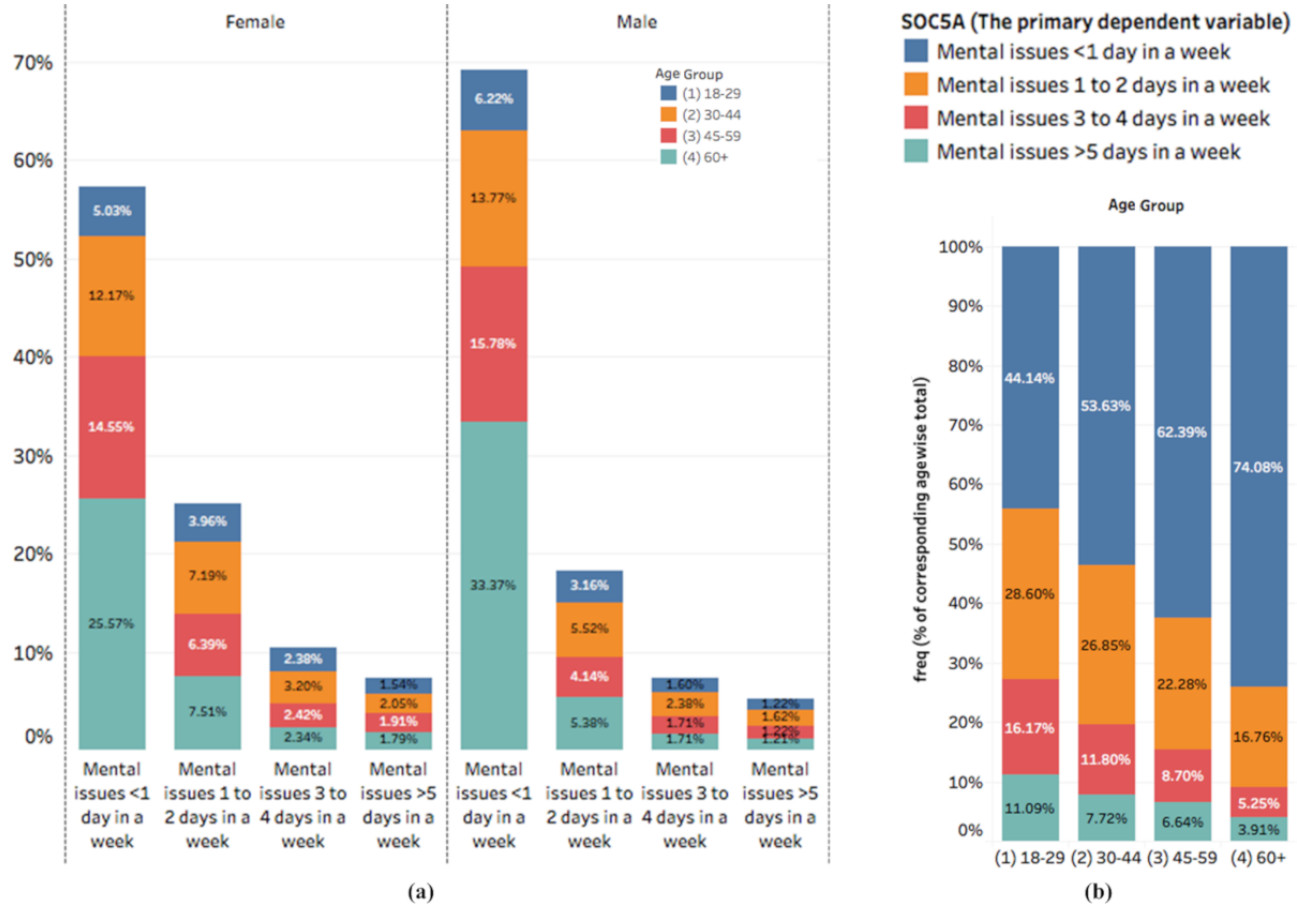
Attribute *soc5a* (felt nervous, anxious, or on edge), attribute *soc5b* (felt depressed), attribute *soc5c* (felt lonely), and attribute *soc5d* (felt hopeless about the future) achieved a Cronbach alpha approximating .8 (Figure 2b), thus confirming their internal consistency and suitability for modeling.

Gender- and Age-Related Variation in Mental Health Indicators

Gender- and age-specific difference was observed in attribute *soc5a*, with females having a higher incidence than males (two proportion *z* test, *P*<.001; Figure 3a) and young adults in the

18-29 years age group having higher incidence than other age groups (P<.001; Figure 3b). The age group 18-29 years in both genders was most vulnerable to mental stress for more than 5 days in a week, thus indicating that COVID-19 may have disproportionately affected the mental health of youth due to a variety of factors.

Figure 3. (a) Genderwise and (b) agewise distribution of mental issues (attribute *soc5a*) variable. Significance was tested using two proportion z test and chi-square test, respectively, showing a higher prevalence of mental issues among youth and in women.



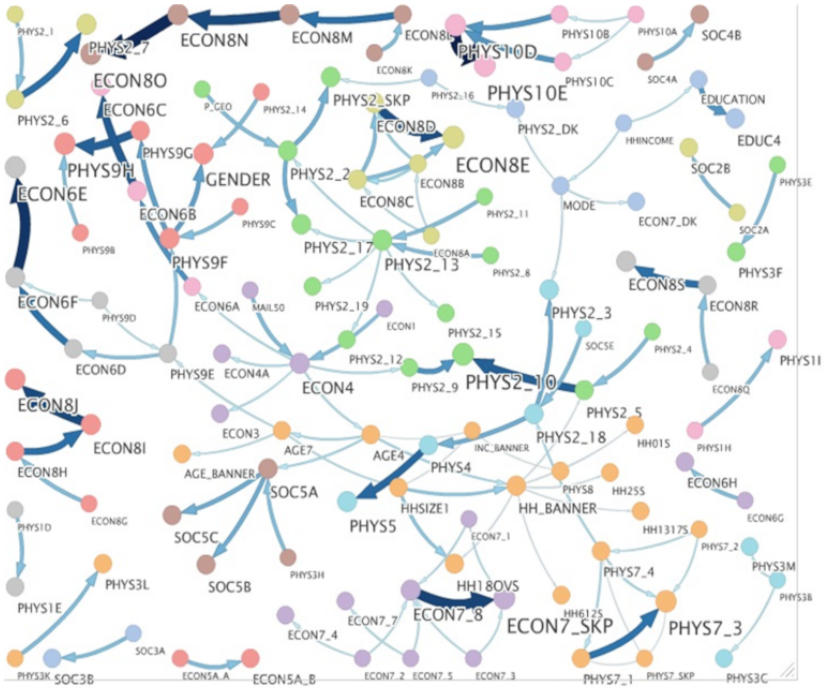
Associations of Anxiety in the United States

A chi-square test revealed many significant associations of the mental health variables (Multimedia Appendix 2). However, this analysis does not account for potential confounding or *explaining away* effects.

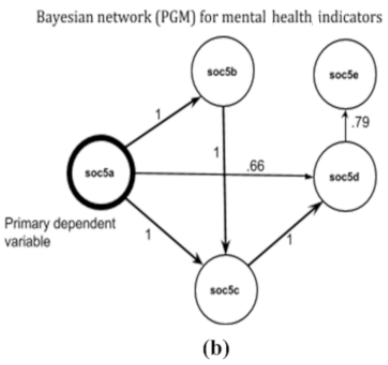
Data-Driven Bayesian Network Analysis

Hence, a data-driven BN structure learning exercise was carried out and revealed interesting findings. From the learned structure, attribute *soc5a* (felt nervous, anxious, or on edge in the last 7 days) was found to be the parent variable for other mental health indicators in almost 100% of the bootstrapped networks, represented as the strength of the edges (Figure 4b). Being a driver variable in the structure, attribute *soc5a* was taken as the primary dependent variable for downstream modeling analysis.

Figure 4. (a) Consensus structure learned through 101 bootstrapped samples. Hill climbing search along with Bayesian information criterion was used to learn the structures and connections having edge strength and direction strength more than 90% are shown. The color of the edges represents the proportion of networks in which that edge was present in the 101 bootstrapped samples, an indicator of confidence; (b) attribute *soc5a* was found to be the parent node of all other mental health variables, therefore, leading to our choice of this variable as the primary dependent variable. PGM: probabilistic graphical model.



(a)



(b)

Impact of Social Life and Work-Related Stressors

Our analysis using network inference via the exact inference algorithm showed a clear impact of in-person social communication on the reduction of anxiety levels. A strong (>5% with CI ~1% on both sides) and (>6.5% with CI ~1.5% on both sides) monotonic increase between control of anxiety and frequency of speaking with neighbors (attribute *soc2a*, attribute *soc2b*) were observed. This effect was weaker (~1.5% with a wide confidence interval) with digital communication with friends and family conducted over phone, text, email, or other internet media (attribute *soc3a*, attribute *soc3b*). This finding underscores the importance of social communication while maintaining the appropriate measures such as masks and social distancing to maintain mental health during such isolating times. We also observed that the presence of kids in the house reduces the probability of depression by >11% with CI ~2% on both sides. Furthermore, the exact inference upon the network revealed an increase in the conditional probability of anxiety (attribute *soc5a*) arising from canceled or postponed work (>4% with CI ~1.4%), canceled or postponed school (7% with CI ~1.5%), working from home (>5% with CI ~1.3%), and studying from home (>7% with CI ~1.8%). Interestingly, although 83% of all volunteers chose to wear the mask, 77% avoided restaurants, and 83% avoided public and crowded places, these measures were not found to be associated with a significant change in anxiety levels as inferred from our model. These inferences are summarized in Figure 5.

Figure 5. Inferences from the Bayesian network. The difference in inferred probability was calculated after conditioning the independent variables. A positive association implies a mental stress–inducing factor, whereas a negative association implies a mental stress reduction factor. The red circle shows the mean value, with green and blue showing confidence intervals. COPD: chronic obstructive pulmonary disease; SNAP: Supplemental Nutrition Assistance Program; TANF: Temporary Assistance for Needy Families.

Impact of Symptoms and Comorbidities

We also investigated the relationship between mental stress and COVID-19 symptoms indicators. The World Health Organization recommends contacting health service providers if any COVID-19 symptoms (attributes *phys1a* to *phys1q*) are experienced within the last 7 days. Our network did not indicate any significant impact of these responses on mental health (attribute *soc5a*), the conditional probability of which remained unchanged (62.2%) across the responses. Although medical conditions (attributes *phys3a* to *phys3m*) are known to increase the risk of serious illness from COVID-19, our model showed that having cancer (attribute *phys3k*) and hypertension (attribute *phys3b*) had a reverse impact on anxiety levels. Those with

cancer had approximately 8.3% (with ~2% CI) higher conditional probability of having less than 1 anxiety-ridden day in a week (>7% effect for hypertension with CI ~1.5%). Additionally, cystic fibrosis (attribute *phys3i*) and liver disease (attribute *phys3j*) had wide confidence intervals with nonsignificant differences in mean values (Figure 5).

Impact of Economic Factors

Receiving income assistance through Social Security improved the conditional probability of less than 1 day of anxiety in a week by 10.4% (with CI ~1.5%) as compared with the segment of people who did not apply or receive it. Just applying for income assistance led to a 4% improvement (Figure 5). Supplemental Social Security (~5.5% with CI ~4%) and health insurance (~5% with CI ~2%) also led to similar results.

In addition to this, older adults (>60 years) found health insurance more relaxing than younger people. COVID-19 has also severely affected the financial condition of individuals, which may also lead to mental stress.

Predictive Modeling for Susceptibility to Anxiety Attacks

Our supervised modeling approach used the Markov blanket of the attribute *soc5a* variable, that is age (attribute *age4)*, physical symptoms in the last 7 days (attribute *phys7_4*), staying at home (attribute *phys2_18*), and prior clinical diagnosis of any mental health condition (attribute *phys3h*) as predictors.

The following three prediction scenarios were considered:

1. Mental issues *less than 1 day* in a week (class 1) versus mental issues *more than 1 day* in a week (class 0)
2. Mental issues *less than 1 day* in a week (class 1) versus mental issues *more than 1 day* in a week (class 0)
3. Mental issues *less than 1 day* in a week (class 1) versus mental issues *more than 1 day* in a week (class 0)

RF models achieved the best performance in comparison with SVM, logistic regression, and naive Bayes models on the basis of standard model performance indicators (accuracy, sensitivity, specificity, area under the receiver operating characteristic curve; summarized in Table 1). We observed a decay (accuracy from 0.80 to 0.64; Table 1) in model predictability as we moved from high risk of depression (case 3) to low risk of depression (case 1; Table 1). Such a trend was visible with all four machine learning techniques we used.

Table 1. Model performance indicators of the supervised model for prediction of stress.

Scenarios	Random forest	Support vector machine	Naive Bayes	Logistic regression
Mental issues less than 1 day in a week (class 1) vs mental issues more than 5 days in a week (class 0)				
Accuracy (±CI)	0.80 (0.016)	0.80 (0.016)	0.77 (0.017)	0.77 (0.017)
Sensitivity (±CI)	0.59 (0.063)	0.56 (0.063)	0.59 (0.063)	0.59 (0.063)
Specificity (±CI)	0.82 (0.016)	0.82 (0.016)	0.79 (0.017)	0.78 (0.017)
AUROC[a] (±CI)	0.71 (0.026)	0.69 (0.026)	0.69 (0.025)	0.68 (0.025)
Mental issues less than 1 day in a week (class 1) vs mental issues more than 3 days in a week (class 0)				
Accuracy (±CI)	0.72 (0.018)	0.72 (0.018)	0.74 (0.017)	0.73 (0.018)
Sensitivity (±CI)	0.6 (0.041)	0.6 (0.041)	0.56 (0.041)	0.57 (0.041)
Specificity (±CI)	0.75 (0.018)	0.75 (0.018)	0.78 (0.017)	0.76 (0.018)
AUROC (±CI)	0.68 (0.022)	0.67 (0.022)	0.67 (0.022)	0.67 (0.022)
Mental issues less than 1 day in a week (class 1) vs mental issues more than 1 day in a week (class 0)				
Accuracy (±CI)	0.66 (0.019)	0.66 (0.019)	0.65 (0.019)	0.62 (0.019)
Sensitivity (±CI)	0.48 (0.027)	0.49 (0.027)	0.45 (0.026)	0.61 (0.026)
Specificity (±CI)	0.77 (0.018)	0.76 (0.018)	0.77 (0.018)	0.64 (0.020)
AUROC (±CI)	0.62 (0.019)	0.62 (0.019)	0.61 (0.020)	0.62 (0.018)

[a]AUROC: area under the receiver operating characteristic curve.

Discussion

Mental health is a public health concern. Mood disorders and suicide-related outcomes have increased substantially over the last decade among all age groups and genders [21,22]. The rapid spread of COVID-19 forced governments worldwide to close public gathering places, schools, colleges, restaurants, and industries. Social isolation, digital communication, and working and schooling from home have become the new normal, and many jobs have been lost. Collectively, this has triggered a high level of anxiety, stress, and depression globally. We did not find studies that have used models to not only predict but also explain the subtle effects of life situations on mental health. An explainable probabilistic graphical modeling approach with bootstraps and exact inference allowed us to capture many of these effects in a robust manner. Our study revealed that individuals with a prior diagnosis of any mental illness are the most vulnerable for mental illness during the COVID-19

pandemic, which recommends building national-level policies to regularly track their mental status and treat them accordingly. Most importantly, our results reiterate the economic underpinnings of a collective mental health response. Income assistance via Social Security or Supplemental Social Security had a demonstrable effect on the alleviation of anxiety as inferred from our model, which provides the first scientific evidence, to the best of our knowledge, proving the utility of such efforts. The extent of such measures' effect may be captured in such modeling studies conducted in various parts of the world, with widely varying assistance structures during this time.

Our findings from the United States can also stimulate further cultural and social research in other geographies with similar or different social structures. For example, the effects of in-person communication, as opposed to digital connectedness, may be different in countries where community living and joint families are still commonplace, such as India. Digital connectedness was not as effective as talking to a neighbor, at least in the United States, highlighting that these have fundamentally different influences on mental health and need to be further explored in systematic studies. We conjecture that such differences may arise from the evolutionary mechanisms that have shaped human societies to live and share in close physical connectedness. Such an effect has been previously shown in primates kept in isolation who display depressive symptoms [23,24]. Similarly, parenting and its association with neuropeptide hormones may partially explain [25] our results that the presence of kids reduces anxiety levels. Interestingly, the COVID-19 pandemic has created a unique natural experiment on the collective mental health response of individuals to a health emergency.

The life cycle of such a response may need to be further studied as the world goes through various phases of the pandemic until its resolution. However, our study indicates that the mental health impact is observable within a span of a few months, especially on young individuals. Further research will be needed, ideally in a longitudinal setting, where the same individuals can be surveyed again to understand the dynamics of the collective mental health response.

Our results also highlight that modern technological development in virtual communication is not able to replace natural socializing. Hence, it becomes imperative to design better and more empathetic technological tools that may shape a society and prevent isolation and alienation even while maintaining physical distancing and preventive measures for limiting spread. Personalization and contextualization of such measures will also be important, as our results indicate that persons with previous mental health conditions may be disproportionately affected.

Finally, our results indicate that it may be possible to identify people at the highest risk of developing mental health disturbances. Our model achieved its best performance for those who were most vulnerable (having mental stress more than 5 days in a week) versus least vulnerable (having no stress or less than 1 day of stress in a week). This can help in the segmentation of vulnerable populations such as frontline health care workers and those who are facing disproportionately higher levels of stress during this time.

A key factor in clinical and public health models is transparency and explainability in the face of complex interactions. Mental health variables are expected to have complex dependencies with potential confounding factors, mediation, and intercausal dependency; therefore, we extended our association analysis with data-driven structure learning of a BN. We preferred this approach over black box machine learning and standard statistical modeling for several reasons. Structure learning allows us to discover and model confounding factors transparently, whereas black box machine learning models such as RFs and gradient boosted machines are not well suited for transparent reasoning. Standard statistical approaches make it humanly impossible to model interactions among hundreds of variables. Structure learning allows discovery and dissection of interactions into mediation, confounding, and intercausal effects. The challenge of incorrect learning is addressed by ensembling many BNs (101 in our case) and choosing the ensemble voted structure. Our artificial intelligence (AI) approach has earlier been validated for public health problems [26,27], and this study demonstrates the underexplored potential of such an approach in complex mental health scenarios.

Our study has a few limitations. Establishing causal inference in cross-sectional data is nearly impossible, and we acknowledge the possibility of confounders. However, this was precisely the reason we chose the structure learning approach, as some of the confounding influences can be transparently discovered and explained. The ensemble voted structure over the sufficiently large number of bootstrapped structures is expected to be robust, as a set of 101 BNs was found to be sufficiently large enough for this study to address the challenge of incorrect learning. Our approach is best suited as a probabilistic reasoning model to explain mental health determinants and to make predictions, a useful outcome in COVID-19–induced mental health morbidity. We could not explain why anxiety levels may be lower in persons with pre-existing cancer or hypertension. This may be a result of reduced work environment–related stress or more contact with family members at home. However, the current data set is not suited to address this at a finer level of explainability. In addition, we could not comment upon the temporality and persistence of these effects. Our results are currently limited to only one geography (ie, the United States). However, the relatively large sample size and multiethnic involvement in the survey makes the model representative for most of the ethnicities and influences across the United States; hence, it is likely to hold true in the United States. Finally, we believe that our study contributes to the use of explainable AI to predict mental health at a population level using survey data, hence making it broadly applicable. Survey data sets are notoriously noisy, and our approach achieved a balance between knowledge discovery and a predictive accuracy of 80%, thus establishing a baseline under a novel scenario. Our algorithms can be used as a screening method for identifying individuals who need help, and further studies with additional measurements and features may increase the accuracy of predictions. Therefore, predictive models for screening and assessing the mental health impact of COVID-19 is a crucial step toward proactive

management and prevention of psychiatric comorbidities as populations continue to fight the pandemic.

Authors' Contributions

VK (vibhor@iiitd.ac.in) and TS (tavpriteshsethi@iiitd.ac.in) serve as co-corresponding authors of this article. VK, TS, and IPJ contributed to the study design. IPJ and AK contributed to the data set. IPJ and RA contributed to the data analysis. IPJ, RA, and TS contributed to the paper writing. VK and TS contributed to the paper review.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Variable groups as indicators.
[XLSX File (Microsoft Excel File), 11 KB - mental_v8i4e25097_app1.xlsx]

Multimedia Appendix 2
Supplementary figures.
[PDF File (Adobe PDF File), 239 KB - mental_v8i4e25097_app2.pdf]

References

1. COVID-19: stimulating the economy and employment: as jobs crisis deepens, ILO warns of uncertain and incomplete labour market recovery. International Labour Organization. URL: https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_749398/lang--en/index.htm [accessed 2020-08-06]
2. Mrklas K, Shalaby R, Hrabok M, Gusnowski A, Vuong W, Surood S, et al. Prevalence of perceived stress, anxiety, depression, and obsessive-compulsive symptoms in health care workers and other workers in Alberta during the COVID-19 pandemic: cross-sectional survey. JMIR Ment Health 2020 Sep 25;7(9):e22408 [FREE Full text] [doi: 10.2196/22408] [Medline: 32915764]
3. The Lancet Psychiatry. Mental health and COVID-19: change the conversation. Lancet Psychiatry 2020 Jun;7(6):463 [FREE Full text] [doi: 10.1016/S2215-0366(20)30194-2] [Medline: 32380007]
4. Pereira-Sanchez V, Adiukwu F, El Hayek S, Bytyçi DG, Gonzalez-Diaz JM, Kundadak GK, et al. COVID-19 effect on mental health: patients and workforce. Lancet Psychiatry 2020 Jun;7(6):e29-e30 [FREE Full text] [doi: 10.1016/S2215-0366(20)30153-X] [Medline: 32445691]
5. Chen Q, Liang M, Li Y, Guo J, Fei D, Wang L, et al. Mental health care for medical staff in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e15-e16 [FREE Full text] [doi: 10.1016/S2215-0366(20)30078-X] [Medline: 32085839]
6. Mahase E. Covid-19: mental health consequences of pandemic need urgent research, paper advises. BMJ 2020 Apr 16;369:m1515. [doi: 10.1136/bmj.m1515] [Medline: 32299806]
7. Pfefferbaum B, North CS. Mental health and the Covid-19 pandemic. N Engl J Med 2020 Aug 06;383(6):510-512. [doi: 10.1056/NEJMp2008017] [Medline: 32283003]
8. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. JMIR Ment Health 2016 Nov 23;3(4):e50 [FREE Full text] [doi: 10.2196/mental.5842] [Medline: 27881357]
9. Bruen AJ, Wall A, Haines-Delmont A, Perkins E. Exploring suicidal ideation using an innovative mobile app-strength within me: the usability and acceptability of setting up a trial involving mobile technology and mental health service users. JMIR Ment Health 2020 Sep 28;7(9):e18407 [FREE Full text] [doi: 10.2196/18407] [Medline: 32985995]
10. 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] [Medline: 32298385]
11. COVID Impact Survey. URL: https://www.covid-impact.org [accessed 2020-05-25]
12. About the survey questionnaire. COVID Impact Survey. URL: https://www.covid-impact.org/about-the-survey-questionnaire [accessed 2020-05-25]
13. Revelle W. psych: procedures for psychological, psychometric, and personality research. The Comprehensive R Archive Network. 2020. URL: https://CRAN.R-project.org/package=psych [accessed 2020-06-26]
14. Gámez JA, Mateo JL, Puerta JM. Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. Data Mining Knowledge Discovery 2010 May 11;22(1-2):106-148. [doi: 10.1007/s10618-010-0178-6]
15. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. Psychometrika 1987 Sep;52(3):345-370. [doi: 10.1007/BF02294361]

16. Yedidia JS, Freeman WT, Weiss Y. Generalized belief propagation. NeurIPS Proceedings. 2001. URL: http://papers.nips.cc/paper/1832-generalized-belief-propagation.pdf [accessed 2020-06-28]
17. Sethi T, Maheshwari S. wiseR: a shiny application for end-to-end Bayesian decision network analysis and web-deployment. The Comprehensive R Archive Network. 2018. URL: https://CRAN.R-project.org/package=wiseR [accessed 2020-05-08]
18. Koller D, Friedman N. Probabilistic Graphical Models: Principles and Techniques. Cambridge, MA: MIT Press; 2009.
19. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intelligence Res 2002 Jun 01;16:321-357. [doi: 10.1613/jair.953]
20. scikit-learn. URL: https://scikit-learn.org/stable/ [accessed 2020-06-26]
21. Bilsen J. Suicide and youth: risk factors. Front Psychiatry 2018;9:540. [doi: 10.3389/fpsyt.2018.00540] [Medline: 30425663]
22. Brådvik L. Suicide risk and mental disorders. Int J Environ Res Public Health 2018 Sep 17;15(9):2028 [FREE Full text] [doi: 10.3390/ijerph15092028] [Medline: 30227658]
23. Hennessy MB, McCowan B, Jiang J, Capitanio JP. Depressive-like behavioral response of adult male rhesus monkeys during routine animal husbandry procedure. Front Behav Neurosci 2014;8:309. [doi: 10.3389/fnbeh.2014.00309] [Medline: 25249954]
24. McKinney WT. Primate social isolation: psychiatric implications. Arch Gen Psychiatry 1974 Sep;31(3):422-426. [doi: 10.1001/archpsyc.1974.01760150122018] [Medline: 4415204]
25. Kohl J, Autry AE, Dulac C. The neurobiology of parenting: a neural circuit perspective. Bioessays 2017 Jan;39(1):1-11 [FREE Full text] [doi: 10.1002/bies.201600159] [Medline: 27921311]
26. Sethi T, Mittal A, Maheshwari S, Chugh S. Learning to address health inequality in the United States with a Bayesian decision network. Proc AAAI Conference Artif Intelligence 2019 Jul 17;33:710-717. [doi: 10.1609/aaai.v33i01.3301710]
27. Awasthi R, Patel P, Joshi V, Karkal S, Sethi T. Learning explainable interventions to mitigate HIV transmission in sex workers across five states in India. arXiv. Preprint posted online on November 30, 2020. [FREE Full text]

Abbreviations

BN: Bayesian network
RF: random forest
SVM: support vector machine

Edited by J Torous; submitted 17.10.20; peer-reviewed by HH Rau, I Gabashvili; comments to author 20.11.20; revised version received 10.12.20; accepted 03.02.21; published 20.04.21.

Please cite as:
Jha IP, Awasthi R, Kumar A, Kumar V, Sethi T
Learning the Mental Health Impact of COVID-19 in the United States With Explainable Artificial Intelligence: Observational Study
JMIR Ment Health 2021;8(4):e25097
URL: https://mental.jmir.org/2021/4/e25097
doi:10.2196/25097
PMID:33877051

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org