JMIR Mental Health

Journal Impact Factor (JIF) (2022): 5.2
Volume 7 (2020), Issue 4 ISSN 2368-7959 Editor in Chief: John Torous, MD, MBI

Contents

Editorial

Original Papers

Review

Letter to the Editor

Editorial

Advancing E-Mental Health in Canada: Report From a Multistakeholder Meeting

Gillian Strudwick[1], RN, PhD; Danielle Impey[2], PhD; John Torous[3], MD, MBI; Reinhard Michael Krausz[4], MD, PhD; David Wiljer[5], PhD

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

[2]Mental Health Commission of Canada, Ottawa, ON, Canada

[3]Division of Digital Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[4]Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

[5]University Health Network, Toronto, ON, Canada

Corresponding Author:
Gillian Strudwick, RN, PhD
Centre for Addiction and Mental Health
1001 Queen St W
Toronto, ON, M6J1H4
Canada
Phone: 1 4165358501
Email: gillian.strudwick@camh.ca

Abstract

The need for e-mental health (electronic mental health) services in Canada is significant. The current mental health care delivery models primarily require people to access services in person with a health professional. Given the large number of people requiring mental health care in Canada, this model of care delivery is not sufficient in its current form. E-mental health technologies may offer an important solution to the problem. This topic was discussed in greater depth at the 9th Annual Canadian E-Mental Health Conference held in Toronto, Canada. Themes that emerged from the discussions at the conference include (1) the importance of trust, transparency, human centeredness, and compassion in the development and delivery of digital mental health technologies; (2) an emphasis on equity, diversity, inclusion, and access when implementing e-mental health services; (3) the need to ensure that the mental health workforce is able to engage in a digital way of working; and (4) co-production of e-mental health services among a diverse stakeholder group becoming the standard way of working.

(JMIR Ment Health 2020;7(4):e19360) doi:10.2196/19360

KEYWORDS

mental health; psychiatry; medical informatics; digital health; nursing informatics

Just before North American travel and face-to-face contact was reduced or eliminated by coronavirus disease (COVID-19), more than 250 people gathered in Toronto, Canada, for the 9th Annual Canadian E-Mental Health Conference [1]. In any given year, 20% of Canadians will have experienced a mental health or addiction related issue [2], and by the age of 40 years, approximately half of Canadians have or will have had a mental illness [3]. With a large number of people requiring access to mental health care and the lack of current face-to-face services in Canada to meet this need, the conference was held to share ideas, leading research and best practices for the use of e-mental health (electronic mental health) technologies to improve care. The conference was founded by the University of British Columbia Department of Psychiatry and coproduced this year by the Mental Health Commission of Canada, University of British Columbia Department of Psychiatry, University Health Network, and the Centre for Addiction and Mental Health. In an effort to ensure the relevancy, timeliness, and potential impact of the topics discussed, the audience was made up of a diverse group of stakeholders such as people with lived experience of mental illness, health providers (including peer support workers, caregivers, and health professionals), researchers, students, administrators, system leaders, vendors/technology developers, and policy makers.

Although there is an abundance of digital technologies used in non–mental health contexts, there are only a limited (but growing) number of examples of meaningful digital technology use among mental health populations in the country [4]. The need to quicken the pace of the adoption of digital technologies within mental health service delivery contexts was discussed.

Many strategies were shared. Common themes through all, regardless of technology or tool, were trust, equity, and training. Interestingly, on the heels of the conference, COVID-19 has acted as an unfortunate but much needed accelerator for the scaling and spread of e-mental health technologies [5] such as increasing access to virtual assessments by a physician in Ontario paid for by a temporary billing code of the provincial insurance plan (Ontario Health Insurance Plan) [6]. Watching many of the strategies discussed just weeks before now being adopted locally and worldwide underscores the potential of focusing on trust, equity, and training.

One central theme of numerous discussions was the close attention that needs to be paid to trust, transparency, human centeredness, and compassion in the development and delivery of digital mental health technologies [7]. These concepts are central to the "successful" uptake of these technologies by people with mental illness, especially given the sensitivity of the mental health clinical area. Trust, transparency, human centeredness, and compassion can be understood in several ways. For example, do those using a digital mental health technology trust the information offered digitally? Do they feel the service offered digitally is as good as (or better than) face-to-face service delivery? Do they feel their information is safe and secure enough to share and answer sensitive questions honestly? Do they feel a health professional will do something about their mental health information that is provided digitally? Without being mindful of these important concepts, the risk is that people with mental illness will not use or meaningfully engage with these technologies, and thus, service delivery would have to remain "in person," which we know is already not meeting the current demand for services.

A second theme that permeated all the workshops and discussions was that of equity, diversity, inclusion and access. These topics must be carefully considered and addressed to ensure that some groups are not disproportionally and/or inadvertency disadvantaged when implementing these technologies. While e-mental health technologies may offer some groups improved quality and access to care, there may be others that are not able to take advantage of this form of care due to accessibility, ability, and language among other reasons. This may have the unintended consequence of creating an even larger equity gap between those who can take advantage of e-mental health technologies and those who cannot. For example, there are still numerous areas of the country with limited internet access and therefore offering e-mental health services to these communities would need to be appropriately supported through addressing significant infrastructure issues. To adequately address important equity, diversity, inclusion and access considerations, it will take new knowledge, specific evidence, and cultural humility to build and adopt smarter approaches to mental health care. Diversity, inclusion, equity, and access are the fundamental starting points for expanding reach and broadening perspectives.

A third theme was that the current mental health workforce is likely not ready at this moment to fully engage in a digital way of working. While it will be important that future mental health professionals like psychiatrists, nurses, and social workers have developed digital health competencies in their entry-to-practice level educational programs, there is a large workforce currently practicing that likely do not possess these competencies. Creative ways to support the development of these competencies is certainly needed, as these mental health professionals are currently practicing within the system and have limited time to engage in formal education. As care is increasingly delivered virtually and uses differing care models than that used today, it is also possible that there are roles that do not exist today (eg, digital specialist or navigator) that may be needed to maximize the potential of digital mental health technologies [8]. Careful consideration must be paid to how the current and future workforce can support a system that is increasingly digital.

In moving the abovementioned agenda forward and improving trust, equity, and training, it will be of great importance to involve all relevant stakeholders (people with lived experience, researchers, health providers, developers, etc) and all relevant communities in coproducing and implementing these e-mental health technologies to ensure success. This is important to ensure the appropriate digital mental health services are developed in the first place, which meet the needs of those who will be using them and are easy to use. Methods for "how to" coproduce and engage diverse stakeholder groups are starting to permeate formal academic literature in the mental health space [9,10]. What is now needed is for these methods to become more commonplace within e-mental health technology implementations. Ideally, co-production, co-design, and meaningful engagement of these relevant stakeholders and communities will become a new standard in which we work.

The group of conference organizers and attendees now have the enormous task ahead of accelerating the lessons learned from the conference during a critical time in recent history. This, of course, needs to be done in a thoughtful way while carefully considering issues of equity, diversity, inclusion, and access. There seems no more urgent a time to ensure that e-mental health technologies are used to deliver necessary care.

For additional information about the conference, review the summary report soon to be published by the Mental Health Commission of Canada.

Conflicts of Interest

None declared.

References

1. Mental Health Commission of Canada. Caring in a Digital World: Introducing Disruptive Change to Mental Health Care. 2020 Presented at: 9th Annual E-Mental Health Conference; Mar 5-6, 2020; Toronto, ON URL: https://www.mentalhealthcommission.ca/sites/default/files/downloads/EMH%20Conference%202020%20Program%20FINAL.pdf
2. Smetanin P, Stiff D, Briante C, Adair C, Ahmad S, Khan M. RiskAnalytica, on behalf of the Mental Health Commission of Canada. 2011. The Life and Economic Impact of Major Mental Illness in Canada to 2041 URL: https://www.mentalhealthcommission.ca/sites/default/files/MHCC_Report_Base_Case_FINAL_ENG_0_0.pdf [accessed 2020-03-30]
3. Mental Health Commission of Canada. 2010. Making the Case for Investing in Mental Health in Canada URL: https://www.mentalhealthcommission.ca/sites/default/files/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf [accessed 2020-03-30]
4. Mahajan S, Schellenberg M, Thapliyal A. Mental Commission of Canada. 2014. E-Mental Health in Canada: Transforming the Mental Health System Using Technology URL: https://www.mentalhealthcommission.ca/sites/default/files/MHCC_E-Mental_Health-Briefing_Document_ENG_0.pdf [accessed 2020-03-30]
5. 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]
6. Ontario Ministry of Health and Ministry of Long-Term Care. 2020 Mar 13. Changes to the Schedule of Benefits for Physician Services (Schedule) in Response to COVID-19 Influenza Pandemic URL: http://www.health.gov.on.ca/en/pro/programs/ohip/bulletins/4000/bul4745.pdf [accessed 2020-03-30]
7. Kemp J, Zhang T, Inglis F, Wiljer D, Sockalingam S, Crawford A, et al. Delivery of Compassionate Mental Health Care in a Digital Technology-Driven Age: Scoping Review. J Med Internet Res 2020 Mar 06;22(3):e16263 [FREE Full text] [doi: 10.2196/16263] [Medline: 32141833]
8. Noel VA, Carpenter-Song E, Acquilano SC, Torous J, Drake RE. The technology specialist: a 21st century support role in clinical care. NPJ Digit Med 2019 Jun 26;2(1):61 [FREE Full text] [doi: 10.1038/s41746-019-0137-6] [Medline: 31388565]
9. 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]
10. McGrath P, Wozney L, Bishop A, Curran J, Chorney J, Rathore SS. Mental Health Commission of Canada. 2018. Toolkit for e-Mental Health Implementation URL: https://www.mentalhealthcommission.ca/sites/default/files/2018-09/E_Mental_Health_Implementation_Toolkit_2018_eng.pdf [accessed 2020-04-24]

Abbreviations

e-mental: electronic mental
COVID-19: coronavirus disease

Edited by A Tal, G Eysenbach; submitted 14.04.20; peer-reviewed by MA Bahrami, E Da Silva; comments to author 23.04.20; revised version received 23.04.20; accepted 24.04.20; published 30.04.20.

Please cite as:
Strudwick G, Impey D, Torous J, Krausz RM, Wiljer D
Advancing E-Mental Health in Canada: Report From a Multistakeholder Meeting
JMIR Ment Health 2020;7(4):e19360
URL: http://mental.jmir.org/2020/4/e19360/
doi:10.2196/19360
PMID:32330114

Original Paper

A Web- and Mobile-Based Intervention for Comorbid, Recurrent Depression in Patients With Chronic Back Pain on Sick Leave (Get.Back): Pilot Randomized Controlled Trial on Feasibility, User Satisfaction, and Effectiveness

Sandra Schlicker[1,2], MSc; Harald Baumeister[3], PhD; Claudia Buntrock[1], PhD; Lasse Sander[4], PhD; Sarah Paganini[5], PhD; Jiaxi Lin[6], PhD; Matthias Berking[1], PhD; Dirk Lehr[7], PhD; David Daniel Ebert[1,8], PhD

[1]Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany
[2]Department of Clinical Psychology and Psychotherapy, Philipps-University Marburg, Marburg, Germany
[3]Department of Clinical Psychology and Psychotherapy, Ulm University, Ulm, Germany
[4]Department of Rehabilitationpsychology and Psychotherapy, Albert-Ludwigs-University Freiburg, Freiburg, Germany
[5]Department of Sport and Sport Science, Albert-Ludwigs-University Freiburg, Freiburg, Germany
[6]Department of Psychiatry and Psychotherapy Medical Center, Albert-Ludwigs-University Freiburg, Freiburg, Germany
[7]Health Psychology and Applied Biological Psychology, Leuphana University Lüneburg, Lüneburg, Germany
[8]Faculty of Behavioural and Movement Sciences, Section of Clinical Psychology, Vrije University Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Sandra Schlicker, MSc
Department of Clinical Psychology and Psychotherapy
Friedrich-Alexander-University Erlangen-Nürnberg
Nägelsbachstraße 25a
Erlangen, 91052
Germany
Phone: 49 91318567564
Email: Sandra.Schlicker@fau.de

Abstract

Background: Chronic back pain (CBP) is linked to a higher prevalence and higher occurrence of major depressive disorder (MDD) and can lead to reduced quality of life. Unfortunately, individuals with both CBP and recurrent MDD are underidentified. Utilizing health care insurance data may provide a possibility to better identify this complex population. In addition, internet- and mobile-based interventions might enhance the availability of existing treatments and provide help to those highly burdened individuals.

Objective: This pilot randomized controlled trial investigated the feasibility of recruitment via the health records of a German health insurance company. The study also examined user satisfaction and effectiveness of a 9-week cognitive behavioral therapy and Web- and mobile-based guided self-help intervention Get.Back in CBP patients with recurrent MDD on sick leave compared with a waitlist control condition.

Methods: Health records from a German health insurance company were used to identify and recruit participants (N=76) via invitation letters. Study outcomes were measured using Web-based self-report assessments at baseline, posttreatment (9 weeks), and a 6-month follow-up. The primary outcome was depressive symptom severity (Center for Epidemiological Studies–Depression); secondary outcomes included anxiety (Hamilton Anxiety and Depression Scale), quality of life (Assessment of Quality of Life), pain-related variables (Oswestry Disability Index, Pain Self-Efficacy Questionnaire, and pain intensity), and negative effects (Inventory for the Assessment of Negative Effects of Psychotherapy).

Results: The total enrollment rate with the recruitment strategy used was 1.26% (76/6000). Participants completed 4.8 modules (SD 2.6, range 0-7) of Get.Back. The overall user satisfaction was favorable (mean Client Satisfaction Questionnaire score=24.5, SD 5.2). Covariance analyses showed a small but statistically significant reduction in depressive symptom severity in the intervention group (n=40) at posttreatment compared with the waitlist control group (n=36; $F_{1,76}$=3.62, P=.03; d=0.28, 95% CI −0.17 to 0.74). Similar findings were noted for the reduction of anxiety symptoms ($F_{1,76}$=10.45; P=.001; d=0.14, 95% CI −0.31

to 0.60) at posttreatment. Other secondary outcomes were nonsignificant (.06≤*P*≤.44). At the 6-month follow-up, the difference between the groups with regard to reduction in depressive symptom severity was no longer statistically significant ($F_{1,76}$=1.50, *P*=.11; *d*=0.10, 95% CI −0.34 to 0.46). The between-group difference in anxiety at posttreatment was maintained to follow-up ($F_{1,76}$=2.94, *P*=.04; *d*=0.38, 95% CI −0.07 to 0.83). There were no statistically significant differences across groups regarding other secondary outcomes at the 6-month follow-up (.08≤*P*≤.42).

Conclusions: These results suggest that participants with comorbid depression and CBP on sick leave may benefit from internet- and mobile-based interventions, as exemplified with the positive user satisfaction ratings. The recruitment strategy via health insurance letter invitations appeared feasible, but more research is needed to understand how response rates in untreated individuals with CBP and comorbid depression can be increased.

Trial Registration: German Clinical Trials Register DRKS00010820; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00010820.

(JMIR Ment Health 2020;7(4):e16398) doi:10.2196/16398

KEYWORDS

pilot project; low back pain; depressive disorder; mental health; sick leave

Introduction

Background

Chronic back pain (CBP) is a pervasive condition with a 12-month prevalence rate of 38% and a lifetime prevalence of approximately 40% in adults [1]. It is also associated with a 2- to 3-fold increased risk for major depressive disorder (MDD) [2], increased morbidity, and diminished quality of life [3,4]. In addition, depression is a core predictor of persistent pain symptoms, increased pain-related disability, and poor treatment outcomes [5-7]. MDD and CBP each account for 2% of disability-adjusted life years worldwide [8], with immense health care and socioeconomic costs due to productivity losses [9]. Thus, from an individual and societal perspective, it is imperative to provide treatment options that decrease patients' burden and specifically target individuals' ability to return to work following sick leave [10].

Effective psychological face-to-face (f2f) treatments exist for depression and CBP [11]. A recent meta-analysis found evidence for the effectiveness of f2f treatments on depression symptoms compared with a nonactive control group (*g*=0.71, 95% CI 0.66 to 0.77) [12]. However, we found no evidence for the effectiveness of multidisciplinary treatments for CBP and comorbid depression. Despite the availability of effective f2f treatments, CBP patients with recurrent depression on sick leave are a difficult-to-reach population with traditional therapy because of a lack of medical and/or disease-related disability specialists.

Internet- and mobile-based interventions (IMIs) have the potential to reach this population because they are easily accessible at any time and in any location. IMIs may be particularly beneficial in psychological and medical treatments as they are accessible and scalable [13]. In addition, the effectiveness of IMIs with mental disorders (eg, depression) [14], disease-related distress in chronic somatic conditions [15], cancer [16], pain [17-19], and coexisting somatic and mental problems (eg, diabetes and depression) [20] is well established.

However, only a few studies have been conducted on the effectiveness of IMIs for individuals with CBP and depression. Recent studies are considering the effectiveness of an IMI on depression in CBP patients following orthopedic rehabilitation, compared with treatment-as-usual (TAU) [21,22]. Irrespective of the findings of these studies, not all patients with CBP seek inpatient rehabilitation treatment. Hence, future research must consider other recruitment strategies. Using health record data might be a valid and innovative recruitment strategy to identify CBP patients with depression.

Objectives

Thus, one aim of this pilot randomized controlled trial (RCT) was to investigate the feasibility of this recruitment strategy as well as the feasibility, user satisfaction, and effectiveness of a guided IMI for CBP patients with depression on sick leave. The program is conceptualized as a stand-alone intervention to provide help to this difficult-to-reach population and complement conventional health care for CBP patients with depression. We expected the IMI to be more effective in reducing depressive symptom severity and pain-associated measures and in increasing the quality of life compared with a waitlist control condition.

Methods

Study Design

This study compares the effectiveness of a guided depression intervention for patients suffering from CBP, resulting in current sick leave, with a waitlist control group (WLC). The intervention was evaluated in a two-armed RCT. The study procedures were approved by the ethical board of the Friedrich-Alexander-University Erlangen-Nürnberg (323_15B), and the trial was registered in the German Clinical Trials Register (DRKS00010820). All study outcomes except for the Structured Clinical Interview (SCID) for the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [23] were measured using Web-based self-report assessments at baseline (t1), posttreatment (t2), and a 6-month follow-up (t3). A secure Web-based system (advanced encryption standard, 256-bit encrypted) was used. This study was initially planned with a target sample of 250 participants. However, the trial did not reach the targeted sample of participants (N=76) due to changes

in personnel in the insurance company responsible for sending invitation letters. Thus, the planned number of invitation letters to be sent (12,000) was not achieved. The study was initially powered to detect medium effect sizes (*d*=0.40; N=200, power of 95%) and accounted for 25% dropout (N=250). Post hoc analysis with N=76 revealed that we were able to detect an effect size *d* of 0.65 with a power of 80%.

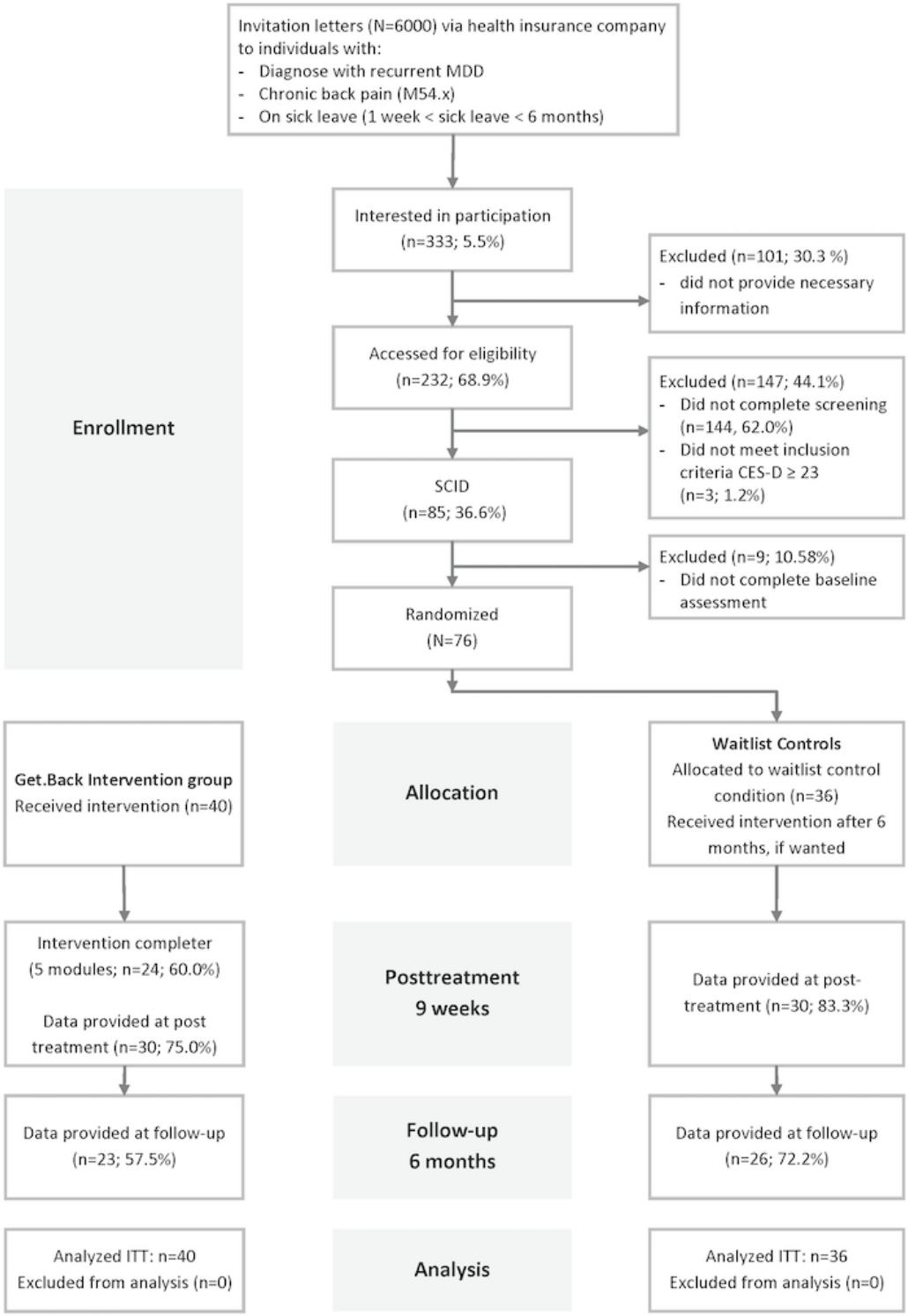
Procedure

Recruitment was carried out by the study team and supported by a German health insurance company (BARMER) from October 2016 until the end of December 2017 by sending invitation letters to policy holders (N=6000). The inclusion criteria were as follows: (1) recurrent diagnosis of MDD and CBP (M54.x according to ICD-10) [24] in the past 16 months, (2) sick leave for more than a week but less than 6 months, (3) no lifetime diagnosis of psychosis, (4) no nursing care level 2 or higher (eg, needing help at least three or more times a day with body care, food, mobility, and household care), and (5) no acute/recent cancer diagnosis in the past 16 months.

In addition, participants were eligible for the study if they (1) were at least 18 years old, (2) had at least moderate depressive symptoms (Center for Epidemiological Studies Depression Scale; CES-D≥23) [25,26], (3) sufficient German language proficiency, and (4) had access to a computer with internet, an email address, and a mobile phone.

The exclusion criteria included the following: (1) current psychotherapeutic treatment, (2) exposure to other online trainings provided by the health insurance company, (3) problems with sight or hearing, and (4) a notable suicidal risk indicated by a score greater than 2 on the Beck Depression Inventory–II item number 9 [27,28] and/or suicidal behavior within the last 5 years (assessed during the SCID) [29]. Individuals who were interested in the study contacted the research team and were asked to fill out a brief online screening form to ensure inclusion criteria were fulfilled. After eligible individuals gave the required informed consent for participation, an account for each participant was created. The account's user name was the participant-provided email address. The account was password protected (password was chosen by participants to prevent misuse of their data). Furthermore, the use of the intervention was free of charge for the study participants. Study procedures are documented in Figure 1.

Figure 1. Flowchart for Get.Back. CES-D: Center for Epidemiological Studies Depression; ITT: intention-to-treat; MDD: major depressive disorder; SCID: Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders; WAI: working alliance.



Randomization and Blinding

Participants eligible for the study were randomly allocated to one of two groups (intervention group [IG] or WLC) based on an a priori defined list after completing the baseline assessment. An automated, Web-based randomization program [30] was used, which features permuted block randomization. Variable randomly arranged block sizes of 4, 6, 8 and an allocation ratio of 1:1 were adopted. An independent research team member not otherwise involved in the study conducted the randomization. Participants were not blinded to treatment condition.

Interventions

All participants had unrestricted access to TAU (eg, visiting a general practitioner). Health care utilization data were collected with the well-validated Trimbos and iMTA Questionnaire for costs associated with psychiatric illness (TiC-P; see outcome measures) [31,32].

Intervention Group

Participants in the IG had access to the intervention Get.Back. Get.Back is adapted from eSano BackCare-D to suit people on current sick leave [21]. The online intervention is based on cognitive behavioral therapy (CBT) and consists of 7 weekly modules lasting 45 to 60 min each. Modules include information regarding psychoeducation, behavioral activation, problem solving, cognitive restructuring, return to work, self-esteem, and relapse prevention (for a detailed description see Lin et al [21]). eSano BackCare-D was originally adapted from GET.ON Mood Enhancer [33,34]. GET.ON Mood Enhancer was proven to be effective in different populations including individuals with MDD alone [34], individuals with MDD and comorbid diabetes [20], and a subclinically depressed population [35-37]. Get.Back differs from eSano BackCare-D mainly because of content regarding returning to work (for detailed information, see Table 1). In eSano BackCare-D, return to work was included as an optional module, whereas in Get.Back, this module was integrated into the obligatory modules and was extended and improved in content. This module specifically provides stress management strategies (coping with solvable and unsolvable problems in the workplace), psychoeducational information on how to adapt the workplace to each individual's needs (eg, ergonomic chair and desk arrangement), and relaxation and exercise information to facilitate motion and prevent pain. The return to work module was introduced in the fifth intervention module. The optional modules on partnership, sexuality, and sleep habits from eSano BackCare-D were also used as optional modules in Get.Back. In addition to eSano BackCare-D, we also included 4 optional minimodules (15 min each) on perfectionism, social support, communication, and appreciation that could be completed after module 3, 4, 5, or 6, respectively (for more detailed information, see Table 2). These topics play an important role in acclimating to the workplace after sick leave, and thus, it is crucial to address such information. We also included 1 booster module 4 weeks after the completion of the intervention contrary to 2 booster modules in eSano BackCare-D. The emphasis is on homework assignments, which ideally leads to the application of the learned skills into daily routines. Interactive elements (eg, emails and text messages), reminders, and exercises were used to enhance adherence to the intervention (for detailed information about the intervention, see Multimedia Appendix 1).

Table 1. Content of the Get.Back intervention and changes from eSano BackCare-D.

Modules[a]	Depression-specific topics	Back pain–specific topics
1	• Psychoeducation	• Psychoeducation
2	• Behavioral activation	• Pain-related complications
3	• Problem solving	• Problem solving
4	• Cognitive restructuring	• Pain-related rumination
5	• My way back to work: Stress management strategies, psychoeducational information on personal needs at the workplace, relaxation and exercises to facilitate motion and prevent pain, and coping with pain in a work-related environment[b]	• My way back to work: Stress management strategies, psychoeducational information on personal needs at the workplace, relaxation and exercises to facilitate motion and prevent pain, and coping with pain in a work-related environment[b]
6	• Mood and self esteem[b] • Fostering exercises to value oneself[b]	• Strengths and successes despite pain
7	• Relapse prevention	• Building up and maintaining resources
8	• Booster session (within 4 weeks after the regular modules)	• Booster session (within 4 weeks after the regular modules)

[a]Original intervention: eSano BackCare-D [21].

[b]Adaptations made to the original intervention content for the Get.Back intervention.

Table 2. Content of the optional and minimodules in the Get.Back intervention.

Module	Topics
Minimodules (15 min)[a]	On perfectionism, social support, communication, and appreciation
Perfectionism	Information and exercises on how to cope with perfectionism (especially related to the work environment), cognitive restructuring for a more relaxed and tension free attitude towards tasks
Social Support	Information and exercises on how to receive and provide social support if needed, interaction in difficult situations (work-related conflicts)
Communication	Introduction to a concept of nonviolent communication and exercises to facilitate interaction with colleagues and supervisors
Appreciation	Introduction of mindfulness-based ideas and exercises on how to appreciate positive aspects in daily life routine
Optional modules (45-60 min)	Healthy sleep & intimacy and partnership

[a]Adaptations made to the original intervention content for the Get.Back intervention.

Intervention Guidance

Participants were guided by trained psychologists, called eCoaches, who provided semistandardized feedback within 2 working days after each completed module. The feedback was based on an eCoach manual, which is intended to ensure adherence to the treatment. The manual also includes instructions to remind, set deadlines, and formulate standardized feedback. The communication between eCoaches and participants occurred through Get.Back's online platform. The feedback content was based on the participant's statements and included positive reinforcement to encourage participants to continue with the training. If any further questions arose, participants and eCoaches were able to contact each other at any time via the platform. In case of noncompletion of the modules, eCoaches sent reminders to participants. eCoaches received a training based on the eCoach manual and on previous experiences by the trainers as well as constant supervision during their time as eCoaches on this study. Training and supervision were provided by a trained and fully licensed (according to German laws and regulations) behavioral and cognitive psychotherapist.

Text Message Coach

Participants had the option to receive daily standardized text messages to increase treatment outcomes and adherence, as well as to support transferring learned skills into their daily routine. Content included the following: (1) reminders to complete weekly assignments, (2) repetition of the content, and (3) motivation enhancement components. Each participant received a total of 42 text messages.

Waitlist Control Group

Participants in the WLC had access to the unguided intervention after study completion in addition to unrestricted access to TAU throughout their participation.

Outcomes

Primary Outcome - Depressive Symptom Severity at Posttreatment

Depressive symptom severity was measured with CES-D [25,26], a widely used instrument in IMI depression trials [14,38]. The 20 items refer to the previous week and are answered on a 4-point Likert scale, ranging from 0 (*rarely or none of the time*) to 3 (*most or all of the time*) with a total score ranging from 0 to 60. Items include the most common symptoms related to depression, such as low mood, loss of appetite, concentration difficulties, and hopelessness. CES-D scores of 16 or greater indicate clinically relevant levels of depression severity. The CES-D has been shown to have excellent reliability (ie, internal consistency of Cronbach alpha=.89) [39]. In this study, Cronbach alpha was .82.

Secondary Outcomes

Depression Symptoms

Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR16) [40,41] is a 16-item questionnaire that assesses all criteria for MDD according to DSM-5 [42]. The items refer to the previous week and are answered on a 4-point Likert scale, ranging from 0 (absence of symptom for 7 days) to 3 (presence of intense symptoms every day). Total scores range from 0 to 27, with the following cutoffs: 0 to 5 indicates no depression, 6 to 10 indicates mild depression, 11 to 15 indicates moderate depression, 16 to 20 indicates severe depression, and 21 to 27 indicates very severe depression. Psychometric properties are reported to be adequate (Cronbach alpha was .77) [43]. In this study, Cronbach alpha was .74.

Quality of Life

Assessment of Quality of Life (AQoL-6D) [44] was used to measure the health-related quality of life. AQoL-6D includes 20 items and covers 6 dimensions. Psychometric properties of AQoL-6D are well established [44]. In this study, Cronbach alpha was .85. We also used EuroQoL [45,46], a widely implemented instrument, which covers 5 health domains. In this study, Cronbach alpha was .76.

Anxiety

The Hamilton Anxiety and Depression Scale [47,48] is a 7-item self-report measure that assesses anxiety and depressive symptoms during the last 7 days on two subscales. In this study, only the anxiety subscale was used. Items are answered on a 4-point Likert scale with total scores ranging from 0 to 21 on the anxiety subscale. Cutoffs are as follows: 8 to 10 indicates mild anxiety, 11 to 14 indicates moderate anxiety, and 15 to 21 indicates severe anxiety. Psychometric properties are reported to be adequate [47]. Cronbach alpha in this study was .70.

Pain-Related Disability

The Oswestry Disability Index [49,50] is a 10-item self-report questionnaire with good validity and reliability [51]. Total scores can also be used to calculate a correlated functional disability at the individual level (measured in percentages, ranging from 0% to 100%). Cutoffs are as follows: 0% to 20% indicates minimal disability, 21% to 40% indicates moderate disability, 41% to 60% indicates severe disability, 61% to 80% indicates crippled, and 81% to 100% indicates individuals who are either bedbound or exaggerating their symptoms [49,52]. In this study, Cronbach alpha was .90.

Pain Rating

We used three items measured on an 11-point numerical (0-10) scale regarding the worst, least, and average pain during the last week. The three items were averaged to calculate a global pain rating over the past week. In addition, we assessed pain using a categorical rating of pain intensity (none, mild, moderate, and severe).

Pain-Related Self-Efficacy

The Pain Self-Efficacy Questionnaire [53,54] is a valid and reliable 10-item instrument that assesses self-efficacy expectations related to pain on a 7-point Likert scale. Total scores range from 0 to 60, and higher scores represent more self-efficacy. In this study, Cronbach alpha was .89.

Screening for Bipolar Disorder

The Mood Disorder Questionnaire (MDQ) [55] is a brief self-report instrument which comprises three sections. In the first section, 13 manic and hypomanic symptoms are assessed on a dichotomous scale (ie, yes and no). Section two asks if any of these symptoms are experienced at the same time, which is also answered on a dichotomous scale. Section three is answered on a 4-point Likert scale (no problems to serious problems) regarding the degree to which their symptoms have caused problems. The screening is considered positive if a cutoff of ≥7 symptoms for section one, *yes* in section two, and a problem severity of *moderate* or *serious* in section three are indicated. Psychometric properties are well validated, with a reported sensitivity of 0.28 and a specificity of 0.97 [56].

Working Capacity

The Subjective Prognostic Employment Scale [57] is a 3-item self-report questionnaire, with a sum score from 0 to 3. It is well validated (internal consistency according to Guttman scale: rep=0.99) [57]. The rep=0.99 refers to the coefficient of reproducibility and can be considered as a measure of internal consistency. The coefficient ranges from 0 (no reproducibility of data) to 1 (perfect reproducibility of data) with values of 0.90 and above indicating acceptable reproducibility.

Client Satisfaction

The Client Satisfaction Questionnaire (CSQ) [58,59] adapted for the assessment of client satisfaction in IMIs by Boß et al [60] consists of 8 items that are rated on 4-point and 5-point Likert scales. CSQ was only assessed in the IG. In this study, Cronbach alpha was .93.

Adverse Events

The Inventory for the Assessment of Negative Effects of Psychotherapy (INEP) [61] was used to assess negative effects during posttreatment online assessments. The 15-item INEP assesses common changes participants may have experienced in line with the intervention's 5 domains (intrapersonal change, relationship, friends and family, work, and stigma). This study had a Cronbach alpha of .55. To further assess serious adverse events (SAE), participants were asked about adverse events at the beginning of each module and were also encouraged to report any such events to their eCoach who monitored the SAEs and initiated further actions if needed. Symptom deterioration was assessed by calculating the reliable change index [62] for CES-D (for a detailed description, see Statistical Analysis).

Working Alliance

To evaluate a subjective rating of the alliance between eCoach and patient, the short, revised version of the working alliance (WAI) [63,64] was administered only after the third module (half of the intervention). The WAI–short revised is a well validated [65], 12-item questionnaire and consists of three subscales assessing (1) how closely the client and therapist agree on and are mutually engaged in the goals of treatment (task subscale); (2) how closely the client and therapist agree on how to reach the treatment goals (goal subscale); and (3) the degree of mutual trust, acceptance, and confidence between the client and therapist (bond subscale). Items are rated on a 5-point Likert scale from 1 (*seldom*) to 5 (*always*). Cronbach alpha in this study was .92.

Health Care Utilization and Sick Leave Data

Health care utilization data and data on sick leave were collected with the well-validated TiC-P illness [31,32] via online self-report.

Adherence

The attrition rate was calculated by identifying the percentage of individuals who no longer utilized the intervention, as indicated in their log-in data. This provides an estimate of the participants' intervention adherence.

Statistical Analysis

All analyses were conducted using SPSS Statistics Version 25 (IBM Corporation) [66] and are reported in accordance with the Consolidated Standards of Reporting Trials statement [67]. Missing data were multiply imputed using a Markov chain Monte Carlo [68] multivariate imputation algorithm with 50 estimations per missing value in accordance with the intention-to-treat principle. Descriptive statistics were reported for feasibility of recruitment, intervention usage, client satisfaction, and relationship with the eCoach. Analyses of covariance adjusted for sex, age, and baseline symptom severity were performed to analyze primary and secondary outcomes between groups at posttreatment and the 6-month follow-up. In a sensitivity analysis, the same analyses were performed with the last observation carried forward (LOCF) method for the postassessment and follow-up. In addition, we performed per-protocol analyses to assess differences in the primary outcome between intervention completers and noncompleters.

Participants were classified as intervention completers if they adhered to at least 80% of the intervention (5 out of 7 modules).

Results are reported as mean within- and between-group differences and as Cohen *d* effect sizes (and their 95% CIs, according to Hedges and Olkin [69]) controlling for baseline scores (ie, calculating change scores divided by the pooled standard deviation of change scores). To assess improvements in the primary outcome (depressive symptom severity) at the individual level, treatment response and near-to-symptom-free status (eg, CES-D<16) were calculated at posttreatment and the 6-month follow-up. In addition, corresponding numbers needed to treat (NNT, with 95% CI) to achieve symptom-free status were calculated at posttreatment and the 6-month follow-up. Treatment response was defined as a 50% symptom reduction from baseline to follow-up, as well as based on the reliable change index by Jacobson and Truax [62]. Participants with a reliable positive change in depression (RCI>1.96; CES-D≥ −12.10; CES-D points take into account the reliability of the CES-D to compensate for measurement errors) were classified as responders to the intervention. Accordingly, symptom deterioration was classified as an increase in 7.8 CES-D points between baseline and posttreatment assessments, and between baseline and the 6-month follow-up. Statistical significance in all analyses was set at alpha<.05 and was one-sided according to Cho and Abe [70].

Results

Descriptive Statistics

In total, 76 participants were included in the study. For detailed information on characteristics, see Table 3. There were no clinically relevant differences in baseline characteristics between the groups.

The posttreatment (9-week) questionnaire return rate was 79% (60/76). Of those, 75% (30/40) of participants were in the IG and 83% (30/36) of participants were in the WLC. Complete data at the 6-month follow-up were collected from 58% (23/40) of participants in the IG and 72% (26/36) of participants in the WLC, with an overall completion rate of 64% (49/76). Dropout rates did not statistically differ at posttreatment (X^2_1=0.7, *P*=.37) or at the 6-month follow-up (X^2_1=1.7, *P*=.18). Participants in the study were predominately female with an average age of 50.78 years (SD 7.85). The majority of participants had a midlevel of education (equivalent to General Educational Development Test) and were married. The average age at depression onset was 35.19 years (SD 14.64), and the average number of previous depressive episodes was 8.2 (SD 7.27). Self-reported depressive symptom severity measured with CES-D was 32.92 (SD 7.52). The most common depressive episode severity (QIDS) was moderate (24/76, 32%) or severe (28/76, 37%). In total, 9% (7/76) of participants screened positive for bipolar disorder (MDQ). The pain-related disability (ODI) was 27.3%, which corresponds to a moderate disability. The average pain intensity was 4.39 (SD 1.94, range 0-11), corresponding to a moderate level of pain present during the last week. The most common categorical rating on the actual pain intensity was moderate (39/76, 51%).

Table 3. Demographics and clinical characteristics.

Variable	IG[a] (n=40)	WLC[b] (n=36)	Total (N=76)
Age (years), mean (SD)	51.3 (8.60)	50.1 (7.00)	50.78 (7.85)
Sex, female, n (%)	26 (65)	29 (81)	55 (72)
Education, n (%)			
High	8 (20)	9 (25)	17 (22)
Medium	26 (65)	25 (69)	51 (67)
Low	6 (15)	2 (6)	8 (11)
Marital status[c], n (%)			
Single/separated	8 (20)	14 (39)	22 (29)
Married/in a relationship	31 (78)	21 (58)	52 (68)
Widowed	1 (3)	1 (3)	2 (3)
Number of depressive episodes, mean (SD)	7.85 (6.39)	8.60 (8.26)	8.20 (7.27)
Age at onset (years), mean (SD)	36.4 (14.4)	33.8 (14.9)	35.19 (14.6)
Severity of current episode[d], n (%)			
Mild	9 (23)	5 (14)	14 (18)
Moderate	12 (30)	12 (33)	24 (32)
Severe	15 (38)	13 (36)	28 (37)
Very severe	4 (10)	6 (17)	10 (13)
Current pain intensity, n (%)			
None	3 (8)	4 (11)	7 (9)
Mild	15 (38)	12 (33)	27 (36)
Moderate	20 (50)	19 (53)	39 (51)
Severe	2 (5)	1 (3)	3 (4)
Social support[c], n (%)			
High	10 (25)	10 (28)	20 (26)
Medium	12 (30)	14 (39)	26 (34)
Low	18 (45)	12 (33)	30 (39)
Partial disability, yes, n (%)	0 (0)	3 (8)	3 (4)
Positive screening for bipolar disorder, yes, n (%)	4 (10)	3 (8)	7 (9)

[a]IG: intervention group.

[b]WLC: waitlist control group.

[c]Percentages less than 100 are due to missing data.

[d]Measured with Quick Inventory of Depressive Symptomatology.

Use of Other Health Care Services and Sick Leave Change

Data on concurrent mental health care service use was provided by 58% (44/76) of participants at the 6-month follow-up (IG: 21/44, 48%; WLC: 23/44, 52%). In total, 84% (37/44) of participants reported visits to their GP in the previous 3 months, with more participants (21/23, 91%) in the WLC compared with the IG (16/21, 76%; X^2_1=1.8, *P*=.17). About one-third of participants reported visits to a psychotherapist and/or a specialist in neurology and psychiatry, with no notable differences between study groups (15/44, 34%; IG: 7/21, 33%; WLC: 8/23, 35%; X^2_1=0.0, *P*=.83). Approximately two-thirds of participants (26/40, 65%), with an equal number of participants in the IG (13/26, 50%) and WLC (13/26, 50%), used pain management medication for back pain with no difference between the groups (X^2_1=0.0, *P*>.99). Half of the participants (23/39, 59%) took antidepressant medication (IG: 12/23, 52%; WLC: 11/23, 48%) with no statistical difference between the groups (X^2_1=0.2, *P*=.60). Data on current sick leave were provided by 50% of participants (IG: 16/40, WLC: 22/36). At the 6-month follow-up, 33% (25/76) of study participants reported being on sick leave during the last 3 months. There

were more participants in the WLC (n=17/22) reporting to have been on sick leave than in the IG (n=8/16). However, there was no statistically significant difference between the groups (X^2_1=3.0, *P*=.08).

Feasibility: Feasibility of Recruitment, Intervention Usage, Client Satisfaction, and Relationship With the eCoach

Feasibility of Recruitment

Of the 6000 individuals who were sent invitations, interest in the study was expressed by 333 (5.50%) individuals. However, only 3.86% (232/6000) of individuals started the screening process. Of those 232 individuals, 144 (62.0%) did not complete the screening, while 3 participants (1.2%) did not meet the inclusion criteria. In total, 36.6% (85/232) of the screened individuals were eligible for study participation and were invited for a diagnostic interview via telephone (see Figure 1). Of these 85 individuals, 9 (11%) did not complete the baseline assessment after the telephone interview and were excluded, resulting in 76 (N) study participants. In total, the enrollment rate of those who received invitation letters was 1.26% (76/6000).

In terms of costs, the total cost of recruitment was 2683.20€ (US $2974.13), and corresponding costs of approximately 8.05€ (US $8.92) per individual signing up for participation. The cost associated with every finally enrolled individual (ie, intervention implementation costs) was 35.30€ (US $39.13) per person.

Intervention Usage

Participants completed on average 4.8 (SD 2.6) modules of the intervention. In total, 60% (24/40) of participants in the IG were identified as completers, and 55% (22/40) of participants adhered to all 7 modules. Of the 16 (16/40, 40%) participants who did not complete at least 5 modules, 1 (3%) participant never started the intervention. Completers and noncompleters did not differ in their baseline characteristics.

Client Satisfaction

Participants were generally satisfied with the intervention. The average score on the CSQ-8 was 24.53 (SD 5.20, range 8-32, min=10, max=32). A high quality and satisfaction rating of the intervention was reported by 90% (26/29) of participants, who stated that they would recommend the intervention to a friend. The vast majority of participants (25/29, 86%) stated that they would use the intervention again if the need arose. Four-fifths of the participants received the training that they wanted (24/29, 83%), perceived the intervention as helpful in dealing with their problems more effectively, and were overall satisfied with the treatment (23/29, 79%). Three-quarters of participants (22/29) also reported that the intervention met their needs and that they were satisfied with the amount of help they received throughout the intervention.

Relationship With the eCoach

Analysis of WAI-SR showed a good WAI between participants and eCoaches with a mean score of 39.30 (SD 11.64, range 15-60, min=21, max=56). Participant's ratings of the subscales revealed the highest ratings in the subscale task (mean 14.22, SD 3.77, range 5-20, min=8, max=20), followed by the goal subscale (mean 13.94, SD 3.70, range 5-20, min=8, max=20) and the bond subscale (mean 11.13, SD 4.98, range 5-20, min=4, max=19).

Short-Term Effects

Primary Intervention Outcome

The mean scores for outcomes are reported in Table 4. Table 5 displays results for all outcome measures. The results revealed statistically significant reductions in the primary outcome from baseline to posttreatment in both the IG (reduction of 6.84 points on CES-D; t_{40}=5.82, *P*<.001; *d*=0.84, 95% CI 0.39 to 1.30) and WLC (reduction of 4.64 points on CES-D; t_{36}=3.86, *P*<.001; *d*=0.64, 95% CI 0.17 to 1.12). There was a statistically significant difference between the IG and WLC at posttreatment, resulting in a small between-group effect size favoring the intervention condition ($F_{1,76}$=3.62, *P*=.03; *d*=0.28, 95% CI −0.17 to 0.74). There were no significant differences in the primary outcome between intervention completers and noncompleters ($F_{1,29}$=0.01; *P*=.97).

Table 4. Mean (SD) of outcomes.

Outcomes	Baseline			Posttreatment			6-month follow-up		
	IG[a] (n=40), mean (SD)	WLC[b] (n=36), mean (SD)	Total (N=76), mean (SD)	IG (n=40), mean (SD)	WLC (n=36), mean (SD)	Total (N=76), mean (SD)	IG (n=40), mean (SD)	WLC (n=36), mean (SD)	Total (N=76), mean (SD)
CES-D[c]	32.50 (7.27)	33.55 (7.84)	32.92 (7.52)	25.66 (8.48)	28.91 (6.38)	27.20 (7.68)	24.36 (9.03)	26.40 (7.13)	25.37 (8.20)
QIDS[d]	15.00 (4.57)	15.55 (4.53)	15.62 (4.53)	13.06 (4.35)	14.21 (3.15)	13.60 (3.85)	12.76 (4.32)	14.25 (3.54)	13.46 (4.08)
AQoL-6D[e]	0.48 (0.16)	0.47 (0.16)	0.48 (0.16)	0.55 (0.17)	0.51 (0.14)	0.53 (0.15)	0.60 (0.18)	0.55 (0.12)	0.57 (0.16)
EQ-5D-5L[f]	0.64 (0.21)	0.66 (0.18)	0.65 (0.19)	0.67 (0.19)	0.68 (0.17)	0.67 (0.18)	0.69 (0.17)	0.68 (0.15)	0.68 (0.16)
HADS[g]-anxiety	12.18 (3.47)	11.80 (3.30)	12.00 (3.37)	9.34 (3.43)	11.20 (3.11)	10.22 (3.39)	8.57 (3.21)	9.56 (3.21)	9.04 (3.23)
ODI-fd[h]	28.50 (17.97)	26.11 (16.79)	27.39 (17.35)	28.26 (16.29)	25.56 (16.52)	26.98 (16.53)	25.15 (13.43)	24.90 (15.27)	25.03 (14.23)
Average pain intensity	4.68 (1.94)	4.08 (1.91)	4.39 (1.94)	4.68 (1.86)	3.81 (1.76)	4.27 (1.85)	3.89 (1.60)	3.67 (1.80)	3.79 (1.69)
PSEQ[i]	33.72 (12.30)	34.75 (11.50)	34.21 (11.86)	36.67 (11.87)	36.62 (9.38)	36.65 (10.69)	40.14 (13.42)	38.20 (9.70)	39.22 (11.77)
SPE[j]	0.95 (0.74)	0.99 (0.82)	0.97 (0.78)	1.03 (0.66)	0.9 (0.76)	0.97 (0.71)	1.69 (0.34)	1.6 (0.42)	1.65 (0.38)

[a]IG: intervention group.

[b]WLC: waitlist control group.

[c]CES-D: Center of Epidemiological Studies Depression Scale.

[d]QIDS: Quick Inventory of Depressive Symptomatology.

[e]AQoL-6D: Assessment of Quality of life.

[f]EQ-5D-5L: EuroQol.

[g]HADS: Hamilton Anxiety and Depression Scale.

[h]ODI-fd: Oswestry Disability Index-functional disability, measured as % (SD).

[i]PSEQ: Pain Self-Efficacy Questionnaire.

[j]SPE: Subjective Prognosis of Employment Scale.

Table 5. Results of all outcomes.

Time points and outcomes	*F* test (*df*)	*P* value	Between-group		IG[a]		WLC[b]	
			d	95% CI	*d*	95% CI	*d*	95% CI
Posttreatment								
Primary outcome								
CES-D[c]	3.62 (1,76)	0.03	0.28	−0.17 to 0.74	0.86	0.39 to 1.30	0.64	0.17 to 1.12
Secondary outcomes								
QIDS[d]	1.24 (1,76)	0.13	0.16	−0.28 to 0.62	0.43	0.01 to 0.88	0.34	−0.12 to 0.82
AQoL-6D[e]	0.99 (1,76)	0.16	0.2	−0.24 to 0.66	0.39	−0.04 to 0.84	0.26	−0.19 to 0.73
EQ-5D-5L[f]	0.01 (1,76)	0.44	0.07	−0.37 to 0.52	0.14	−0.29 to 0.58	0.09	−0.36 to 0.55
HADS[g]-anxiety	10.45 (1,76)	0.001	0.14	−0.30 to 0.60	0.81	0.36 to 1.27	0.18	−0.27 to 0.65
ODI[h]	0.15 (1,76)	0.35	0.02	−0.42 to 0.47	0.01	−0.42 to 0.45	0.03	−0.42 to 0.49
Average pain intensity	3.76 (1,76)	0.06	0.23	−0.22 to 0.68	0	−0.44 to 0.43	0.14	−0.32 to 0.60
PSEQ[i]	0.02 (1,76)	0.43	0.11	−0.33 to 0.56	0.24	−0.19 to 0.68	0.17	−0.28 to 0.64
SPE[j]	1.35 (1,76)	0.12	0.24	−0.20 to 0.70	0.11	−0.32 to 0.56	0.11	−0.34 to 0.58
6-month follow-up								
Primary outcome								
CES-D	1.50 (1,76)	0.11	0.1	−0.34 to 0.46	0.98	0.51 to 1.46	0.94	0.45 to 1.43
Secondary outcomes								
QIDS	1.93 (1,76)	0.08	0.23	−0.21 to 0.69	0.5	0.06 to 0.95	0.32	−0.14 to 0.79
AQoL-6D	1.44 (1,76)	0.11	0.21	−0.23 to 0.66	0.65	0.20 to 1.10	0.55	0.08 to 1.02
EQ-5D-5L	0.06 (1,76)	0.38	0.16	−0.29 to 0.61	0.22	−0.21 to 0.66	0.08	−0.37 to 0.54
HADS-anxiety	2.94 (1,76)	0.04	0.38	−0.07 to 0.83	1.07	0.60 to 1.54	0.68	0.21 to 1.16
ODI	0.11 (1,76)	0.36	0.14	−0.30 to 0.59	0.21	−0.22 to 0.65	0.07	−0.38 to 0.53
Average pain intensity	0.03 (1,76)	0.42	0.21	−0.24 to 0.66	0.44	0.00 to 0.88	0.22	−0.24 to 0.68
PSEQ	0.57 (1,76)	0.22	0.23	−0.21 to 0.68	0.39	−0.04 to 0.83	0.51	0.04 to 0.98
SPE	0.96 (1,76)	0.16	0.15	−0.29 to 0.61	−1.27	0.79 to 1.76	−0.91	0.43 to 1.41

[a]IG: intervention group.

[b]WLC: waitlist control group.

[c]CES-D: Center of Epidemiological Studies Depression Scale.

[d]QIDS: Quick Inventory of Depressive Symptomatology.

[e]AQoL-6D: Assessment of Quality of Life.

[f]EQ-5D-5L: EuroQol.

[g]HADS: Hamilton Anxiety and Depression Scale.

[h]ODI: Oswestry Disability Index.

[i]PSEQ: Pain Self-Efficacy Questionnaire.

[j]SPE: Subjective Prognosis of Employment Scale.

Treatment Response

Reliable change did not significantly differ between participants in the IG (17/40, 43%) and WLC (11/40, 31%; X^2_1=1.1, *P*=.14; NNT=8, 95% CI 3 to 10^6). A nonsignificant score reduction of 50% from baseline to posttreatment was seen more often in the IG (2/77, 3%) compared with the WLC (n=0; X^2_1=1.8, *P*=.08; NNT=20, 95% CI 9 to 10^6).

Near-to-Symptom-Free Status

Significantly more participants in the IG (5/40, 13%) reached a symptom-free status compared with the WLC (n=0; X^2_1=4.8, *P*=.01; NNT=8, 95% CI 5 to 45).

Secondary Outcomes

The IG showed a significantly greater reduction in anxiety compared with the WLC ($F_{1,76}$=10.45, *P*=.001; *d*=0.14, 95%

CI −0.31 to 0.60) with a within-group effect size *d* of 0.81 (95% CI 0.36 to 1.28; t_{40}=5.40; *P*<.001) versus 0.18 (95% CI −0.27 to 0.65; t_{36}=1.26; *P*=.21) in the WLC. There were no statistically significant differences between the IG and WLC with regard to any other secondary outcomes (eg, pain-related disability, self-reported depressive symptoms, pain-related self-efficacy, quality of life, and subjective prognosis of employment; see Table 5).

Adverse Events

At posttreatment, 10% (4/40) of participants reported at least 1 negative event related to the intervention. In total, 6 negative events were reported by the IG, with the most commonly reported negative event being: "Since the start of the intervention, I suffer more from events in the past" (n=3). In addition, 17% (5/30) of participants reported at least 1 negative event not related to the training. Symptom deterioration did not take place in the IG. In the WLC, 3% (1/37) of participants did experience deterioration. This difference was not statistically significant (X^2_1=1.1; *P*=.14).

Long-Term Effects

Primary Intervention Outcome

Both study groups displayed statistically significant reductions in depressive symptom severity from baseline to the 6-month follow-up (IG: t_{40}=5.99, *P*<.001; *d*=0.98; 95% CI 0.51 to 1.46 and WLC: t_{36}=4.99, *P*<.001; *d*=0.94; 95% CI 0.43 to 1.45); however, the between-group difference was not statistically significant ($F_{1,76}$=1.50, *P*=.11; *d*=0.10, 95% CI −0.34 to 0.46).

Treatment Response, Near-to-Symptom-Free Status, and Symptom Deterioration

A reliable change from baseline to the 6-month follow-up was more often seen in the IG (9/40, 23%) compared with the WLC (6/36, 17%). However, this difference was not statistically significant (X^2_1=0.4, *P*=.52; NNT=17, 95% CI 5 to 10^6). A symptom reduction of 50% from baseline to follow-up was seen in twice as many participants in the IG (6/40, 15%) compared with the WLC (3/36, 8%), but this difference was not statistically significant (X^2_1=0.8; *P*=.18). In all, 48% (19/40) of participants in the IG and 39% (14/36) of participants in the WLC reached symptom-free status at the 6-month follow-up, with no statistically significant difference between the groups (X^2_1=0.5; *P*=.22). From baseline to follow-up, symptom deterioration occurred more often in the WLC, with 6% (2/36) of participants, compared with 3% (2/40) of participants in the IG; however, this difference was not statistically significant (X^2_1=0.4; *P*=.24).

Secondary Outcomes

Analyses revealed that the between-group difference in anxiety at posttreatment was also statistically significant at follow-up ($F_{1,76}$=2.94, *P*=.047; *d*=0.38, 95% CI −0.07 to 0.83). There were no statistically significant differences across groups with regard to any other secondary outcomes (eg, pain-related disability, self-rated depressive symptoms, average pain intensity, pain-related self-efficacy, quality of life, or subjective prognosis of employment; Table 5).

Sensitivity Analysis

Results of the sensitivity analyses were similar to the results of the main analyses. It has previously been shown that LOCF estimates similar effect sizes, but overestimates the precision, compared with multiple imputation [71]. Results of the sensitivity analyses are presented in Table 6.

Table 6. Sensitivity analyses (last observation carried forward).

Time points and outcomes	*F* test (*df*)	*P* value	Between-group		IG[a] (n=40)		WLC[b] (n=36)	
			d	95% CI	*d*	95% CI	*d*	95% CI
Posttreatment								
Primary outcome								
CES-D[c]	3.64 (1,76)	0.03	0.34	−0.11 to 0.80	0.61	0.16 to 1.06	0.34	−0.12 to 0.81
Secondary outcomes								
QIDS[d]	3.79 (1,76)	0.02	0.37	−0.08 to 0.83	0.39	−0.04 to 0.84	0.17	−0.29 to 0.64
AQoL-6D[e]	2.60 (1,76)	0.05	0.37	−0.08 to 0.82	0.34	−0.10 to 0.78	0.11	−0.35 to 0.58
EQ-5D-5L[f]	1.94 (1,76)	0.08	0.35	−0.10 to 0.81	0.23	−0.20 to 0.68	0	−0.46 to 0.46
HADS-anxiety[g]	9.34 (1,76)	0	0.75	0.28 to 1.22	0.62	0.18 to 1.08	0.07	−0.39 to 0.54
ODI[h]	2.38 (1,76)	0.06	0.37	−0.08 to 0.83	0.11	−0.32 to 0.56	0.05	−0.41 to 0.52
Average pain intensity	1.24 (1,76)	0.13	0.08	−0.37 to 0.53	0.04	−0.39 to 0.48	0.08	−0.37 to 0.55
PSEQ[i]	3.62 (1,76)	0.03	0.46	0.01 to 0.92	0.23	−0.21 to 0.67	0.04	−0.42 to 0.50
SPE[j]	0.25 (1.76)	0.3	0.13	−0.32 to 0.58	0.1	−0.34 to 0.54	0.03	−0.43 to 0.49
6-month follow-up								
Primary outcome								
CES-D	3.35 (1,76)	0.04	0.31	−0.14 to 0.77	0.69	0.24 to 1.14	0.4	−0.06 to 0.87
Secondary outcomes								
QIDS	3.81 (1,76)	0.02	0.42	−0.03 to 0.88	0.44	0.00 to 0.89	0.16	−0.30 to 0.63
AQoL-6D	2.36 (1,76)	0.06	0.34	−0.11 to 0.80	0.55	0.10 to 1.00	0.31	−0.15 to 0.78
EQ-5D-5L	2.56 (1,76)	0.06	0.44	−0.01 to 0.90	0.37	−0.07 to 0.82	0	−0.46 to 0.47
HADS-anxiety	3.95 (1,76)	0.03	0.51	0.05 to 0.97	0.71	0.26 to 1.17	0.31	−0.15 to 0.78
ODI	4.89 (1,76)	0.02	0.53	0.08 to 0.99	0.26	−0.18 to 0.70	0.08	−0.38 to 0.55
Average pain intensity	0.73 (1,76)	0.19	0.3	−0.15 to 0.75	0.32	−0.12 to 0.77	0.09	−0.36 to 0.56
PSEQ	3.13 (1,76)	0.04	0.42	−0.03 to 0.88	0.4	−0.04 to 0.84	0.03	−0.42 to 0.50
SPE	0.00 (1,76)	0.98	0.03	−0.41 to 0.49	0.63	0.18 to 1.08	0.58	0.11 to 1.05

[a]IG: intervention group.

[b]WLC: waitlist control group.

[c]CES-D: Center of Epidemiological Studies Depression Scale.

[d]QIDS: Quick Inventory of Depressive Symptomatology.

[e]AQoL-6D: Assessment of Quality of Life.

[f]EQ-5D-5L: EuroQol.

[g]HADS: Hamilton Anxiety and Depression Scale.

[h]ODI: Oswestry Disability Index.

[i]PSEQ: Pain Self-Efficacy Questionnaire.

[j]SPE: subjective Prognosis of Employment Scale.

Discussion

Principal Findings

Delivery of CBT via the internet seems feasible in a highly burdened sample, and the enrollment rate was 1.26% (76/6000). As hypothesized, Get.Back demonstrated small but statistically significant effects compared with the WLC in terms of reducing depressive symptom severity at posttreatment. However, findings did not support effectiveness with regard to pain measures, quality of life, or long-term effectiveness.

Comparison With Previous Research

To the best of our knowledge, there are no published studies regarding digital or f2f psychological interventions for patients with comorbid depression and CBP on sick leave. Our findings regarding the feasibility and user satisfaction of IMIs are in line

with other studies for monodisorder depression and for comorbid depression with somatic diseases [19,72].

The within-group effect size *d* of 0.86 in favor of the IG is comparable with existing evidence for digital interventions for MDD. Königbauer et al [14] found standardized within-group effect sizes (Hedge *g*) ranging from −0.64 (95% CI −1.27 to −0.01) to −1.52 (95% CI −2.22 to −0.82) for the reduction in depressive symptom severity at posttreatment. However, between-group effect sizes found in this study were smaller than those in similar trials on IMIs with depressed individuals. One reason might be the notable improvements in the WLC. Participants in the WLC knew that they were scheduled to get access to Get.Back after a waiting period. Therefore, there is a possibility of an expectancy effect. A recent study showed that patients with MDD who were scheduled to wait for treatment showed a significant decline in depressive symptoms [73]. However, WLCs may also experience a nocebo effect, such that participation as a waitlist control might reduce natural recovery [74]. Future studies are needed to better understand the effects of scheduled waiting in clinical trials.

Moreover, regardless of the treatment format, psychological interventions for depression in individuals with CBP might achieve lower effects compared with individuals without CBP. This hypothesis is supported by a meta-analysis showing that depression treatments tend to be less effective in individuals with general medical disorders compared with nonmedical populations [12].

There is existing evidence that CBT is effective in reducing depressive symptom severity in individuals with CBP [21]. However, this evidence is limited to samples with unspecified depression diagnoses and symptoms. Moreover, internet-based self-help could be less suitable for this group compared with f2f psychotherapy. To date, there is no RCT in an f2f setting that investigates the effectiveness of a CBT depression intervention in individuals with CBP, and only 1 other trial that investigates iCBT in individuals with CBP and clinical depression [21]. We found no trials focusing on individuals with chronic depression or current sick leave.

Furthermore, this study aimed to reach individuals who are not actively seeking help, meaning that their motivation to change might be lower compared with individuals who do actively seek help. In Germany, the decision to pursue psychological treatments for mental illness is made after an average waiting period of approximately 7 years [75]. Therefore, motivation for change can be considered an important predictor in depression treatments and related outcomes [76]. However, it could be that Get.Back combined with our recruitment strategy resulted in smaller effects compared with recruitment strategies directly targeting individuals who are actively seeking help. Thus, it is possible that this intervention may have increased effects if recruitment strategies actively targeted and increased the motivation for change before starting the intervention.

Yet, the between-group effect size *d* of 0.28 was higher than the minimal important difference defined as a standardized mean difference of 0.24, pinpointing the cutoff of clinical relevance in depression treatment [77]. Thus, Get.Back may be a promising treatment for this burdened population. To investigate the beneficial effects of Get.Back, a larger trial with sufficient power is needed on enhancing the overall treatment effect.

Limitations

First, our findings should be interpreted as that of a pilot trial with limited power. Initially, the study was planned as an RCT with a total sample of 250 participants and was designed to specifically explore effects on return to work and cost-effectiveness of Get.Back. Our small sample size reduced the power to detect medium effect sizes. Second, the sample characteristics may have also limited the generalizability of our findings. The percentage of well-educated women was higher than in the general chronic pain population. Third, participants were recruited from a health insurance company. Therefore, results may not be generalizable to other settings in routine, clinical mental health care. Fourth, no clinical interviews took place at posttreatment or the 6-month follow-up. Therefore, changes in the diagnosis of MDD could not be analyzed. Future trials should therefore investigate the potential beneficial effects of the intervention with an extended follow-up period and with sufficient power. Fifth, the WAI version that was used in this study was not adapted for the use of internet interventions. Hence, exploring the agreement on goals might be difficult, as the goals are typically set by the intervention. An adapted version of WAI [78] for use in the internet interventions has been released and should be used in future studies.

Implications for Clinical Practice and Recommendation for Future Research

Our findings have several implications for clinical practice. First, the results of this study suggest that a combined psychological treatment for patients on sick leave with comorbid recurrent depression and CBP might be beneficial. Available treatments generally only focus on one condition, rarely on both. Results of our study show that combining treatments for both conditions within one IMI appears to be feasible with high user satisfaction and acceptable adherence. However, although we found significant effects with respect to the primary outcome, the intervention was not found to be superior with regard to a range of secondary outcomes. It is unclear whether this finding is a result of the low power in this study or due to minimal efficacy for psychological interventions targeted at individuals with comorbid depression and CBP on sick leave. Hence, future studies are needed to compare different treatment modalities (eg, IMI vs f2f) with regard to both effectiveness and reach in the target group, given the challenging sample characteristics.

Second, using health insurance data to address individuals with CBP and a history of depression appears to be a promising strategy to reach individuals in need of treatment, as shown by the initial response rate of 5.5%. However, the rate of actual enrollment (76/6000, 1.26%) is lower than the response rates found in studies aimed at reducing mild to moderate depression and absenteeism in individuals at high risk for taking depression-related sick leave [73]. The requirement that individuals *opt in* to a study based on a postal invitation may have compounded difficulties of recruiting participants with depression.

Moreover, reacting to an invitation letter (eg, completing a baseline assessment and providing informed consent) may have been too demanding for individuals with severe depression. Individuals with depression might be interested in participating in internet-based interventions but not in a clinical trial. This may be particularly pertinent for individuals with CBP and comorbid depression. Consequently, future research should implement measures to reduce participant burden.

Another strategy to further enhance the potential of recruitment via a health insurance company could be the implementation of acceptance facilitation interventions (AFIs). The effectiveness of AFIs has been evaluated in recent research [79-81]. Such interventions may aim to increase the utilization of treatments by directly addressing potential barriers (eg, low outcome expectancy and fear of stigma). Future studies should therefore focus on improving initial response rates to health care insurance letters in addition to increasing conversion rates following expressed interest.

Third, the total cost (2683.20€; US $2973.79) [82], cost for initial response (8.05€; US $ 8.92) [82], and cost per included participant (35.30€; US $39.12) [82] were low. Cost per included participant was comparable with studies using Facebook ads (US $51.70; 46.64€) [82,83]. Compared with the high cost associated with non- or delayed treatments for multimorbid patients with chronic disease and depression, this cost is negligible. A meta-analysis concluded that it is difficult to assess the overall effectiveness of any particular recruitment strategy as some strategies that work well for a certain population may not be optimal for another population; they also discussed the necessity of additional research to better understand effective recruitment strategies [84]. For our studied population, the current recruitment strategy via health insurance letter invitations appeared feasible, but more research is needed to understand how response rates in untreated individuals with CBP and comorbid depression can be increased.

Conclusions

To the best of our knowledge, this is one of the first RCTs investigating the effects of a psychological intervention in individuals with comorbid depression and CBP on sick leave. Despite our inability to examine the actual effects on return to work rates and cost-effectiveness of Get.Back, this trial shows that this particular group of individuals may benefit from IMIs, as shown by the positive user satisfaction ratings. However, besides larger follow-up confirmatory trials, future studies should implement strategies that could better reach the target sample, test possibilities to increase intervention effects, and identify subgroups of patients that may or may not benefit from such interventions and could otherwise be referred to other treatment modalities.

Acknowledgments

The authors acknowledge the support by Deutsche Forschungsgemeinschaft and Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) within the funding program Open Access Publishing. They also acknowledge the support of the BARMER health care insurance company.

This study was carried out with in-house funds of FAU and supported by a grant from the German Federal Ministry of Education and Research (project “Effectiveness of a guided Web-based intervention for depression in back pain rehabilitation aftercare,” grant numbers: 01GY1330A; 01GY1330B). The funders had no role in the study design, data collection and analysis, the decision to publish, or the preparation of the manuscript. The views expressed are those of the authors and not necessarily those of the funders.

Individual participant data are available on request after deidentification beginning 12 months following the article publication. Data will be made available to researchers who provide a methodologically sound proposal, not already covered by others. Proposals should be directed to the corresponding author. Data requestors will need to sign a data access agreement. Provision of data is subject to data security regulations. Investigator support depends on available resources.

Authors' Contributions

SaS, HB, and DE initiated and designed the study; BARMER (German health insurance company) supported the recruitment. SaS, HB, JL, SP, LS, DL, MB, and DE adapted the intervention content and assessments. SaS, HB, and DE were responsible for the recruitment. SaS was responsible for the trial management, analyses, and preparing the first draft of this manuscript. CB and DE supervised the writing process. All authors critically revised the manuscript and approved the final draft of this manuscript.

Conflicts of Interest

All authors were involved in the development of Get.Back or its predecessor versions. SaS and LS have received payments for workshops on e-mental-health. SaS has received reimbursement of congress attendance and travel costs, as well as payments for lectures with the Psychotherapy Training Institutes. HB, DL, and MB received consultancy fees, reimbursement of congress attendance, and travel costs, as well as payments for lectures with the Psychotherapy and Psychiatry Associations and Psychotherapy Training Institutes (discussing E-Mental-Health topics). They have been the beneficiaries of study support (third-party funding) from several public funding organizations. DE possess shares in the GET.ON Institut GmbH, which works to transfer research findings on internet- and mobile phone–based health interventions into routine care. DE has received payments from several companies and health insurance providers for advice on the use of internet-based interventions. He has received payments for

lectures delivered for Psychotherapy and Psychiatry Associations and has been the beneficiary of third-party funding from health insurance providers. DL is minor stakeholder of the GET.ON Institut GmbH, which aims to transfer scientific knowledge related to this research into routine health care. MB, HB, DE, and DL were not involved in the data analysis.

Multimedia Appendix 1
Overview of the Get.Back intervention, summary of content of each module, and screenshots of the intervention.
[PDF File (Adobe PDF File), 1306 KB - mental_v7i4e16398_app1.pdf]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V1.6.1).
[PDF File (Adobe PDF File), 3612 KB - mental_v7i4e16398_app2.pdf]

References

1. Manchikanti L, Singh V, Falco FJ, Benyamin RM, Hirsch JA. Epidemiology of low back pain in adults. Neuromodulation 2014 Oct;17(Suppl 2):3-10. [doi: 10.1111/ner.12018] [Medline: 25395111]
2. Demyttenaere K, Bruffaerts R, Lee S, Posada-Villa J, Kovess V, Angermeyer MC, et al. Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys. Pain 2007 Jun;129(3):332-342. [doi: 10.1016/j.pain.2007.01.022] [Medline: 17350169]
3. Baumeister H, Hutter N, Bengel J, Härter M. Quality of life in medically ill persons with comorbid mental disorders: a systematic review and meta-analysis. Psychother Psychosom 2011;80(5):275-286. [doi: 10.1159/000323404] [Medline: 21646822]
4. Katon WJ. Epidemiology and treatment of depression in patients with chronic medical illness. Dialogues Clin Neurosci 2011;13(1):7-23 [FREE Full text] [Medline: 21485743]
5. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. Arch Intern Med 2003 Nov 10;163(20):2433-2445. [doi: 10.1001/archinte.163.20.2433] [Medline: 14609780]
6. Miles C, Pincus T, Carnes D, Homer KE, Taylor SJ, Bremner SA, et al. Can we identify how programmes aimed at promoting self-management in musculoskeletal pain work and who benefits? A systematic review of sub-group analysis within RCTs. Eur J Pain 2011 Sep;15(8):775.e1-775.11. [doi: 10.1016/j.ejpain.2011.01.016] [Medline: 21354838]
7. Baumeister H, Knecht A, Hutter N. Direct and indirect costs in persons with chronic back pain and comorbid mental disorders--a systematic review. J Psychosom Res 2012 Aug;73(2):79-85. [doi: 10.1016/j.jpsychores.2012.05.008] [Medline: 22789408]
8. World Health Organization. Geneva: World Health Organization; 2018. Global Health Estimates 2016: Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2016 URL: https://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html [accessed 2019-08-17]
9. Statistisches Bundesamt. Statistisches Bundesamt. 2015. Gesundheit: Krankheitskosten URL: https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00033175/2120721159004.pdf%3Bjsessionid%3DB19129ECF3A61F8DB331D847FC4E599C [accessed 2019-08-17]
10. Melkevik O, Clausen T, Pedersen J, Garde AH, Holtermann A, Rugulies R. Comorbid symptoms of depression and musculoskeletal pain and risk of long term sickness absence. BMC Public Health 2018 Aug 6;18(1):981 [FREE Full text] [doi: 10.1186/s12889-018-5740-y] [Medline: 30081870]
11. Guzmán J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary rehabilitation for chronic low back pain: systematic review. Br Med J 2001 Jun 23;322(7301):1511-1516 [FREE Full text] [doi: 10.1136/bmj.322.7301.1511] [Medline: 11420271]
12. Cuijpers P, Karyotaki E, Reijnders M, Huibers MJ. Who benefits from psychotherapies for adult depression? A meta-analytic update of the evidence. Cogn Behav Ther 2018 Mar;47(2):91-106. [doi: 10.1080/16506073.2017.1420098] [Medline: 29345530]
13. 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]
14. Josephine K, Josefine L, Philipp D, David E, Harald B. Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. J Affect Disord 2017 Dec 1;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]
15. Bendig E, Bauereiß N, Ebert D, Snoek F, Andersson G, Baumeister H. Internet- based interventions in chronic somatic disease. Dtsch Arztebl Int 2018 Nov 5;115(40):659-665 [FREE Full text] [doi: 10.3238/arztebl.2018.0659] [Medline: 30381130]
16. Ruland CM, Andersen T, Jeneson A, Moore S, Grimsbø GH, Børøsund E, et al. Effects of an internet support system to assist cancer patients in reducing symptom distress: a randomized controlled trial. Cancer Nurs 2013;36(1):6-17. [doi: 10.1097/NCC.0b013e31824d90d4] [Medline: 22495503]

17. Eccleston C, Fisher E, Craig L, Duggan GB, Rosser BA, Keogh E. Psychological therapies (internet-delivered) for the management of chronic pain in adults. Cochrane Database Syst Rev 2014 Feb 26(2):CD010152 [FREE Full text] [doi: 10.1002/14651858.CD010152.pub2] [Medline: 24574082]
18. Dear BF, Gandy M, Karin E, Staples LG, Johnston L, Fogliati VJ, et al. The Pain Course: a randomised controlled trial examining an internet-delivered pain management program when provided with different levels of clinician support. Pain 2015 Oct;156(10):1920-1935 [FREE Full text] [doi: 10.1097/j.pain.0000000000000251] [Medline: 26039902]
19. Lin J, Paganini S, Sander L, Lüking M, Ebert D, Buhrman M, et al. An internet-based intervention for chronic pain. Dtsch Arztebl Int 2017 Oct 13;114(41):681-688 [FREE Full text] [doi: 10.3238/arztebl.2017.0681] [Medline: 29082858]
20. Nobis S, Lehr D, Ebert DD, Baumeister H, Snoek F, Riper H, et al. Efficacy of a web-based intervention with mobile phone support in treating depressive symptoms in adults with type 1 and type 2 diabetes: a randomized controlled trial. Diabetes Care 2015 May;38(5):776-783. [doi: 10.2337/dc14-1728] [Medline: 25710923]
21. Lin J, Sander L, Paganini S, Schlicker S, Ebert D, Berking M, et al. Effectiveness and cost-effectiveness of a guided internet- and mobile-based depression intervention for individuals with chronic back pain: protocol of a multi-centre randomised controlled trial. BMJ Open 2017 Dec 28;7(12):e015226 [FREE Full text] [doi: 10.1136/bmjopen-2016-015226] [Medline: 29288172]
22. Sander L, Paganini S, Lin J, Schlicker S, Ebert DD, Buntrock C, et al. Effectiveness and cost-effectiveness of a guided internet- and mobile-based intervention for the indicated prevention of major depression in patients with chronic back pain-study protocol of the PROD-BP multicenter pragmatic RCT. BMC Psychiatry 2017 Jan 21;17(1):36 [FREE Full text] [doi: 10.1186/s12888-017-1193-6] [Medline: 28109247]
23. First MB, Gibbon M, Spitzer RL, Williams JB. User's Guide for the Structured Clinical Interview for Dsm-IV Axis I Disorders. New York: American Psychiatric Press; 1997.
24. Dilling H, Mombour W, Schmidt MH. Internationale Klassifikation psychischer Störungen: ICD-10 Kapitel V (F) - Klinisch-diagnostische Leitlinien. Bern: Hogrefe AG; 2014.
25. Hautzinger M, Bailer M. Allgemeine Depressionsskala (ADS): Manual Manual of the CES-D Scale. Göttingen: Beltz Test GmbH; 1993.
26. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
27. Beck AT, Steer A, Brown GK. BDI-II, Beck Depression Inventory: Manual. San Antonio: Psychological Corporation; 1996.
28. Kühner C, Bürger C, Keller F, Hautzinger M. [Reliability and validity of the Revised Beck Depression Inventory (BDI-II). Results from German samples]. Nervenarzt 2007 Jun;78(6):651-656. [doi: 10.1007/s00115-006-2098-7] [Medline: 16832698]
29. Wittchen H, Zaudig M, Fydrich T. Strukturiertes Klinisches Interview für DSM-5 Störungen – Klinische Version. Göttingen, Germany: Hogrefe; 1997.
30. Sealed Envelope. London, UK: Sealed Envelope LTD Create a Blocked Randomisation List URL: https://www.sealedenvelope.com/simple-randomiser/v1/lists [accessed 2019-09-07]
31. Bouwmans C, de Jong K, Timman R, Zijlstra-Vlasveld M, van der Feltz-Cornelis C, Swan ST, et al. Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). BMC Health Serv Res 2013 Jun 15;13:217 [FREE Full text] [doi: 10.1186/1472-6963-13-217] [Medline: 23768141]
32. Hakkaart-van RL, Straten A, Tiemens B, Donker M. Research Gate. Rotterdam, The Netherlands: iMTA; 2002. Manual Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness (TIC-P) URL: https://www.researchgate.net/publication/254758194_Manual_TrimbosiMTA_Questionnaire_for_Costs_Associated_with_Psychiatric_Illness_TIC-P_in_Dutch [accessed 2019-08-17]
33. Ebert DD, Lehr D, Heber E, Riper H, Cuijpers P, Berking M. Internet- and mobile-based stress management for employees with adherence-focused guidance: efficacy and mechanism of change. Scand J Work Environ Health 2016 Sep 1;42(5):382-394 [FREE Full text] [doi: 10.5271/sjweh.3573] [Medline: 27249161]
34. Ebert D, Tarnowski T, Gollwitzer M, Sieland B, Berking M. A transdiagnostic internet-based maintenance treatment enhances the stability of outcome after inpatient cognitive behavioral therapy: a randomized controlled trial. Psychother Psychosom 2013;82(4):246-256. [doi: 10.1159/000345967] [Medline: 23736751]
35. 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. J Am Med Assoc 2016 May 3;315(17):1854-1863. [doi: 10.1001/jama.2016.4326] [Medline: 27139058]
36. Buntrock C, Ebert DD, Lehr D, Cuijpers P, Riper H, Smit F, et al. Evaluating the efficacy and cost-effectiveness of web-based indicated prevention of major depression: design of a randomised controlled trial. BMC Psychiatry 2014 Jan 31;14:25 [FREE Full text] [doi: 10.1186/1471-244X-14-25] [Medline: 24485283]
37. 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]

38. Karyotaki E, Ebert DD, Donkin L, Riper H, Twisk J, Burger S, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clin Psychol Rev 2018 Jul;63:80-92. [doi: 10.1016/j.cpr.2018.06.007] [Medline: 29940401]
39. Stein J, Luppa M. Allgemeine Depressionsskala (ADS). Psychiat Prax 2012;39(6):302-304. [doi: 10.1055/s-0032-1326702]
40. Drieling T, Schärer LO, Langosch J. The Inventory of Depressive Symptomatology: German translation and psychometric validation. Int J Methods Psychiatr Res 2007;16(4):230-236. [doi: 10.1002/mpr.226] [Medline: 18200596]
41. Rush A, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 1;54(5):573-583. [doi: 10.1016/s0006-3223(02)01866-8] [Medline: 12946886]
42. American Psychiatric Association. Diagnostic And Statistical Manual Of Mental Disorders. Washington, USA: American Psychiatric Publishing; 2013.
43. Roniger A, Späth C, Schweiger U, Klein J. A Psychometric Evaluation of the German Version of the Quick Inventory of Depressive Symptomatology (QIDS-SR16) in outpatients with depression. Fortschr Neurol Psychiatr 2015 Dec;83(12):e17-e22. [doi: 10.1055/s-0041-110203] [Medline: 26714254]
44. Richardson JR, Peacock SJ, Hawthorne G, Iezzi A, Elsworth G, Day NA. Construction of the descriptive system for the Assessment of Quality of Life AQoL-6D utility instrument. Health Qual Life Outcomes 2012 Apr 17;10:38 [FREE Full text] [doi: 10.1186/1477-7525-10-38] [Medline: 22507254]
45. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
46. Hinz A, Kohlmann T, Stöbel-Richter Y, Zenger M, Brähler E. The quality of life questionnaire EQ-5D-5L: psychometric properties and normative values for the general German population. Qual Life Res 2014 Mar;23(2):443-447. [doi: 10.1007/s11136-013-0498-2] [Medline: 23921597]
47. Herrmann C, Buss U, Snaith RP. HADS-D - Hospital Anxiety And Depression Scale. Bern: Hans Huber Verlag; 1995.
48. 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]
49. Mannion A, Junge A, Fairbank J, Dvorak J, Grob D. Development of a German version of the Oswestry Disability Index. Part 1: cross-cultural adaptation, reliability, and validity. Eur Spine J 2006 Jan;15(1):55-65 [FREE Full text] [doi: 10.1007/s00586-004-0815-0] [Medline: 15856341]
50. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980 Aug;66(8):271-273. [Medline: 6450426]
51. Wittink H, Turk DC, Carr DB, Sukiennik A, Rogers W. Comparison of the redundancy, reliability, and responsiveness to change among SF-36, Oswestry Disability Index, and Multidimensional Pain Inventory. Clin J Pain 2004;20(3):133-142. [doi: 10.1097/00002508-200405000-00002] [Medline: 15100588]
52. Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine (Phila Pa 1976) 2000 Nov 15;25(22):2940-52; discussion 2952. [doi: 10.1097/00007632-200011150-00017] [Medline: 11074683]
53. Mangels M, Schwarz S, Sohr G, Holme M, Rief W. Der Fragebogen zur Erfassung der schmerzspezifischen Selbstwirksamkeit (FESS). Diagnostica 2009 Apr;55(2):84-93. [doi: 10.1026/0012-1924.55.2.84]
54. Nicholas M. The pain self-efficacy questionnaire: Taking pain into account. Eur J Pain 2007 Feb;11(2):153-163. [doi: 10.1016/j.ejpain.2005.12.008] [Medline: 16446108]
55. Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck PE, et al. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. Am J Psychiatry 2000 Nov;157(11):1873-1875. [doi: 10.1176/appi.ajp.157.11.1873] [Medline: 11058490]
56. Hirschfeld RM, Holzer C, Calabrese JR, Weissman M, Reed M, Davies M, et al. Validity of the mood disorder questionnaire: a general population study. Am J Psychiatry 2003 Jan;160(1):178-180. [doi: 10.1176/appi.ajp.160.1.178] [Medline: 12505821]
57. Mittag O, Raspe H. [A brief scale for measuring subjective prognosis of gainful employment: findings of a study of 4279 statutory pension insurees concerning reliability (Guttman scaling) and validity of the scale]. Rehabilitation (Stuttg) 2003 Jun;42(3):169-174. [doi: 10.1055/s-2003-40095] [Medline: 12813654]
58. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6] [Medline: 10245370]
59. Schmidt J, Lamprecht F, Wittmann WW. [Satisfaction with inpatient management. Development of a questionnaire and initial validity studies]. Psychother Psychosom Med Psychol 1989 Jul;39(7):248-255. [Medline: 2762479]
60. Boß L, Lehr D, Reis D, Vis C, Riper H, Berking M, et al. Reliability and validity of assessing user satisfaction with web-based health interventions. J Med Internet Res 2016 Aug 31;18(8):e234 [FREE Full text] [doi: 10.2196/jmir.5952] [Medline: 27582341]

61. Ladwig I, Rief W, Nestoriuc Y. What are the risks and side effects of psychotherapy? – Development of an inventory for the assessment of negative effects of psychotherapy (INEP). Verhaltenstherapie 2014;24(4):252-263. [doi: 10.1159/000367928]
62. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1037//0022-006x.59.1.12] [Medline: 2002127]
63. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the working alliance inventory. Psychother Res 2006 Jan;16(1):12-25. [doi: 10.1080/10503300500352500]
64. Wilmers F, Munder T, Leonhart R, Herzog T, Plassmann R, Barth J, et al. Die deutschsprachige Version des Working Alliance Inventory – short revised (WAI-SR) – Ein schulenübergreifendes, ökonomisches und empirisch validiertes Instrument zur Erfassung der therapeutischen Allianz. Klin Diagnostik Eval 2008;1(3):343-358 [FREE Full text]
65. Hanson WE, Curry KT, Bandalos DL. Reliability generalization of working alliance inventory scale scores. Educ Psychol Meas 2002;62(4):659-673. [doi: 10.1177/0013164402062004008]
66. IBM. Statistics. J Theoretical Appl Stat 1991;22(3):419-430. [doi: 10.1080/02331889108802322]
67. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001 Apr 14;357(9263):1191-1194. [doi: 10.1016/s0140-6736(00)04337-3] [Medline: 11323066]
68. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002 Jun;7(2):147-177. [doi: 10.1037/1082-989x.7.2.147] [Medline: 12090408]
69. Hedges LV, Olkin I. Statistical methods for meta-analysis. J Educ Stat 1988;13(1):75-78. [doi: 10.2307/1164953]
70. Cho HC, Abe S. Is two-tailed testing for directional research hypotheses tests legitimate? J Bus Res 2013 Sep;66(9):1261-1266. [doi: 10.1016/j.jbusres.2012.02.023]
71. Molenberghs G, Kenward MG. Missing Data In Clinical Studies. Hoboken: John Wiley & Sons; 2007.
72. Reins JA, Boß L, Lehr D, Berking M, Ebert DD. The more I got, the less I need? Efficacy of Internet-based guided self-help compared to online psychoeducation for major depressive disorder. J Affect Disord 2019 Mar 1;246:695-705. [doi: 10.1016/j.jad.2018.12.065] [Medline: 30611913]
73. Ahola P, Joensuu M, Knekt P, Lindfors O, Saarinen P, Tolmunen T, et al. Effects of scheduled waiting for psychotherapy in patients with major depression. J Nerv Ment Dis 2017 Aug;205(8):611-617. [doi: 10.1097/NMD.0000000000000616] [Medline: 27861459]
74. Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, et al. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta Psychiatr Scand 2014 Sep;130(3):181-192. [doi: 10.1111/acps.12275] [Medline: 24697518]
75. Steffanowski A, Löschmann C, Schmidt J, Wittmann W, Nübling R. Metaanalyse Der Effekte Psychosomatischer Rehabilitation. [Meta-Analysis on Effects of Psychosomatic Rehabilitation]. Bern: Huber; 2007.
76. Zuroff DC, Koestner R, Moskowitz DS, McBride C, Marshall M, Bagby MR. Autonomous motivation for therapy: A new common factor in brief treatments for depression. Psychother Res 2007 Mar;17(2):137-147. [doi: 10.1080/10503300600919380]
77. Cuijpers P, Turner EH, Koole SL, van Dijke A, Smit F. What is the threshold for a clinically relevant effect? The case of major depressive disorders. Depress Anxiety 2014 May;31(5):374-378. [doi: 10.1002/da.22249] [Medline: 24677535]
78. Penedo JM, Berger T, Holtforth MG, Krieger T, Schröder J, Hohagen F, et al. The Working Alliance Inventory for guided internet interventions (WAI-I). J Clin Psychol 2019 Jun 25. [doi: 10.1002/jclp.22823] [Medline: 31240727]
79. Ebert DD, Franke M, Kählke F, Küchler AM, Bruffaerts R, Mortier P, WHO World Mental Health - International College Student collaborators. Increasing intentions to use mental health services among university students. Results of a pilot randomized controlled trial within the World Health Organization's World Mental Health International College Student Initiative. Int J Methods Psychiatr Res 2019 Jun;28(2):e1754. [doi: 10.1002/mpr.1754] [Medline: 30456814]
80. Lin J, Faust B, Ebert D, Krämer L, Baumeister H. A web-based acceptance-facilitating intervention for identifying patients' acceptance, uptake, and adherence of internet- and mobile-based pain interventions: randomized controlled trial. J Med Internet Res 2018 Aug 21;20(8):e244 [FREE Full text] [doi: 10.2196/jmir.9925] [Medline: 30131313]
81. Baumeister H, Seifferth H, Lin J, Nowoczin L, Lüking M, Ebert DD. Impact of an acceptance facilitating intervention on patients' acceptance of internet-based pain interventions: a randomized controlled trial. Clin J Pain 2015 Jun;31(6):528-535. [doi: 10.1097/AJP.0000000000000118] [Medline: 24866854]
82. ECB Statistical Data Warehouse. European Central Bank - Statistical Data Warehouse - Quick View URL: https://sdw.ecb.europa.eu/quickview.do;jsessionid=5A9424640335F3F0569D5CFA4D531D4C?SERIES_KEY=120.EXR.D.USD.EUR.SP00.A&start=01-08-2019&end=31-08-2019&submitOptions.x=0&submitOptions.y=0&trans=N [accessed 2019-09-01]
83. Schwinn T, Hopkins J, Schinke SP, Liu X. Using Facebook ads with traditional paper mailings to recruit adolescent girls for a clinical trial. Addict Behav 2017 Feb;65:207-213 [FREE Full text] [doi: 10.1016/j.addbeh.2016.10.011] [Medline: 27835860]
84. Liu Y, Pencheon E, Hunter RM, Moncrieff J, Freemantle N. Recruitment and retention strategies in mental health trials - A systematic review. PLoS One 2018;13(8):e0203127 [FREE Full text] [doi: 10.1371/journal.pone.0203127] [Medline: 30157250]

Abbreviations

AFIs: acceptance facilitation interventions
AQoL-6D: Assessment of Quality of Life
CBP: chronic back pain
CBT: cognitive behavioral therapy
CES-D: Center for Epidemiological Studies Depression Scale
CSQ: Client Satisfaction Questionnaire
DSM: Diagnostic and Statistical Manual of Mental Disorders
f2f: face-to-face
FAU: Friedrich-Alexander-Universität Erlangen-Nürnberg
IMIs: internet- and mobile-based interventions
INEP: Inventory for the Assessment of Negative Effects of Psychotherapy
LOCF: last observation carried forward
MDD: major depressive disorder
MDQ: Mood Disorder Questionnaire
NNT: numbers needed to treat
QIDS-SR16: Quick Inventory of Depressive Symptomatology Self-Report
RCT: randomized controlled trial
SAE: serious adverse events
SCID: Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders
TAU: treatment-as-usual
TiC-P: Trimbos and iMTA Questionnaire for costs associated with psychiatric illness
WAI: working alliance
WLC: waitlist control group

Edited by G Eysenbach; submitted 25.09.19; peer-reviewed by JM Gómez Penedo, K Naversnik; comments to author 18.11.19; revised version received 11.01.20; accepted 27.01.20; published 15.04.20.

Please cite as:
Schlicker S, Baumeister H, Buntrock C, Sander L, Paganini S, Lin J, Berking M, Lehr D, Ebert DD
A Web- and Mobile-Based Intervention for Comorbid, Recurrent Depression in Patients With Chronic Back Pain on Sick Leave (Get.Back): Pilot Randomized Controlled Trial on Feasibility, User Satisfaction, and Effectiveness
JMIR Ment Health 2020;7(4):e16398
URL: http://mental.jmir.org/2020/4/e16398/
doi:10.2196/16398
PMID:32293577

Original Paper

Temporal Associations of Daily Changes in Sleep and Depression Core Symptoms in Patients Suffering From Major Depressive Disorder: Idiographic Time-Series Analysis

Noah Lorenz[1,2], MSc; Christian Sander[1,2], PhD; Galina Ivanova[3], Prof Dr; Ulrich Hegerl[2,4], Prof Dr

[1]Department of Psychiatry and Psychotherapy, Faculty of Medicine, Leipzig University, Leipzig, Germany

[2]Research Centre of the German Depression Foundation, Leipzig, Germany

[3]Institute for Applied Informatics, Leipzig, Germany

[4]Department of Psychiatry, Psychosomatics and Psychotherapy, Goethe-Universität Frankfurt, Frankfurt, Germany

Corresponding Author:
Noah Lorenz, MSc
Research Centre of the German Depression Foundation
Goerdelerring 9
Leipzig, 04109
Germany
Phone: 49 341 2238740
Email: noah.lorenz@medizin.uni-leipzig.de

Abstract

Background: There is a strong link between sleep and major depression; however, the causal relationship remains unclear. In particular, it is unknown whether changes in depression core symptoms precede or follow changes in sleep, and whether a longer or shorter sleep duration is related to improvements of depression core symptoms.

Objective: The aim of this study was to investigate temporal associations between sleep and depression in patients suffering from major depressive disorder using an idiographic research approach.

Methods: Time-series data of daily sleep assessments (time in bed and total sleep time) and self-rated depression core symptoms for an average of 173 days per patient were analyzed in 22 patients diagnosed with recurrent major depressive disorder using a vector autoregression model. Granger causality tests were conducted to test for possible causality. Impulse response analysis and forecast error variance decomposition were performed to quantify the temporal mutual impact of sleep and depression.

Results: Overall, 11 positive and 5 negative associations were identified between time in bed/total sleep time and depression core symptoms. Granger analysis showed that time in bed/total sleep time caused depression core symptoms in 9 associations, whereas this temporal order was reversed for the other 7 associations. Most of the variance (10%) concerning depression core symptoms could be explained by time in bed. Changes in sleep or depressive symptoms of 1 SD had the greatest impact on the other variable in the following 2 to 4 days.

Conclusions: Longer rather than shorter bedtimes were associated with more depression core symptoms. However, the temporal orders of the associations were heterogeneous.

(JMIR Ment Health 2020;7(4):e17071) doi:10.2196/17071

KEYWORDS

depression; sleep; time series; idiographic; self-management

Introduction

Most patients with depression suffer from sleep disturbances, which mainly occur in the form of difficulty in initiating or maintaining sleep. Sleep disturbances have often been reported by patients and clinicians as an early sign or symptom at depression onset. This assumption has been supported by studies showing that insomnia and short sleep duration increase the risk of depression onset [1,2]. In addition, sleep disturbances are a risk factor for recurrent depression [3] and worsening symptoms of depression [4] (also see [5] or [6] for an overview). These studies indicate that sleep plays a causal role in the development of depression. This interpretation was supported by findings of studies focusing on the treatment of insomnia with cognitive

behavior therapy, which has been shown to substantially reduce depression symptoms [7,8]. One study also showed that cognitive behavior therapy focused on treating insomnia symptoms was equally effective in reducing depression core symptoms as therapy focused on treating depression symptoms [9].

However, the association between sleep disturbances and depression can also be explained by a common pathogenic factor that causes both symptoms. The arousal regulation model of affective disorders describes upregulated brain arousal as such a central pathogenic factor [10], which is commonly found in patients with major depressive disorder. The withdrawal and sensation avoidance observed in major depressive disorder is considered to be an autoregulatory reaction of the individual in response to the high brain arousal. Upregulated brain arousal can also provide an explanation for other coexisting symptoms such as insomnia or anxiety. According to this concept, arousal regulation can be influenced by factors such as increasing sleep pressure via sleep deprivation, which counteracts the upregulated brain arousal, whereas extended sleep durations maintain the pathogenic factor (arousal). These assumptions were supported by an experimental study revealing that an increase in time in bed and total sleep time has negative effects on mood [11]. Even partial sleep deprivation of 2 hours was shown to have an antidepressive effect in some patients [12]. Moreover, one of the decisive interventions applied in cognitive behavioral therapy for insomnia is sleep restriction, in which patients experience a significant reduction in bedtime. Thus, consistent with sleep restriction studies in the context of depression, it is possible that this reduction in time in bed leads to an improvement in depression symptoms through the mechanism described by the arousal model. This association has already been convincingly demonstrated for patients with bipolar disorder in a time-series study design [13]. The authors investigated the temporal association between sleep and mood over the course of an average of 169 days, finding that a decrease in sleep duration was followed by an increase in mood the following day for 34% of the patients, and the same association was found between bedrest (awake in bed) and mood for 22% of the patients, with no positive cross-correlations identified for sleep duration. However, this is one of only few studies that have investigated the relationship between sleep and mood in an idiographic research design.

Idiographic research is based on within-subject analysis techniques. Instead of analyzing cohorts to explain variance in the population, the aim of an idiographic approach is to explain variance within individuals without seeking to generalize the results to other individuals. Such analyses require repeated measurements for the same individual. The experience sampling method or momentary ecological assessment is a promising research technique to collect repeated measures within individuals for an idiographic analysis [14]. One of the main advantages of idiographic research is that it is a more person-centered approach, since patterns within individuals can be investigated over time. For example, this method allows for investigating temporal associations with various time lags between several variables for a given individual, enabling identifying indications for cause-effect relationships of several variables within an individual [15]. Therefore, this approach can provide results with high clinical relevance, as they apply to specific individuals in specific contexts. In particular, if interindividual differences between patients are to be assumed, an idiographic research design can avoid the disadvantages of group statistics that risk blurring the relevant effects on the individual level. Molenaar et al [16] provide a more detailed discussion and comparison of interindividual and intraindividual research methodologies.

For both clinicians and individual patients, it would be highly relevant to understand how changes in sleep are related to changes in depressive symptoms. For example, if shorter bedtimes lead to a reduction in depressive symptoms in individual patients, this information can be used by both the clinician for treatment and by the patient for self-management. Idiographic time-series studies are useful to explore these questions, and some authors have shown that ecological momentary assessment in combination with time-series analysis are useful tools for this purpose [17,18].

The goal of the present study was to investigate the relationship between sleep and depression core symptoms by applying selected time-series methods on long-term collected sleep and self-rating data. We pursued three explorative aims. The first aim was to investigate the number of patients demonstrating significant temporal associations (indicating possible causal relationships) between time in bed/total sleep time and core depression symptoms. The second aim was to describe the nature of the relationship from two aspects: whether changes in sleep cause changes in depression or vice versa, and whether more or less sleep causes a reduction in depression. Finally, we sought to analyze the impact and temporal dynamics of the effects in patients with significant temporal associations by calculating the amount of variance explained in the time series and by analyzing how changes of 1 SD in a time series affect the other time series over a 10-day period.

Methods

Sensor-Based System for Therapy Support and Management of Depression Research Project

The data analyzed in this study were obtained from the research project Sensor-based System for Therapy Support and Management of Depression (STEADY), with the broad aim of developing a digital solution allowing patients to collect data on the course of their disease using smartphones and mobile sensors. To support patients in their self-management, one aim of STEADY was to identify patient-specific patterns and associations in the recorded self-ratings and digital behavior markers that can serve as a basis for patient-specific self-management. A feasibility study was carried out within the STEADY project, which was divided into three distinct study phases targeted at investigating specific modules or features of the comprehensive STEADY system. Before and after each study phase, data were collected using printed self-assessment questionnaires, including the Pittsburg Sleep Quality Index (PSQI) [19]. During the study phase, patients were supported by weekly and later monthly telephone calls. Furthermore, a

monthly on-site appointment was held at the research center of the German Depression Foundation in Leipzig, Germany.

Participants

Participants of the STEADY feasibility study were primarily recruited at the Department of Psychiatry and Psychotherapy of the University of Leipzig, Germany. In addition, interested individuals could register in the study via a project-specific website. The study was approved by the local ethics committee of the University of Leipzig. Once written informed consent was obtained, participants were invited for a diagnostic interview to check for eligibility according to the following criteria. Inclusion screening criteria were at least 18 years of age, diagnosis of a recurrent depressive disorder, current depressive symptomatology reflected by a minimum of 14 points on the Inventory of Depressive Symptomatology Clinician (IDS-C) scale [20], undergoing professional treatment with regard to the diagnosis, and living in close proximity to the study center to facilitate regular study appointments. Exclusion criteria were psychiatric comorbidities (alcohol/drug addiction, schizophrenia, schizotypal and delusional disorders, borderline personality disorder), severe somatic disorders, acute suicidal behavior, pregnancy/lactation, and electronic implants.

Inclusion and exclusion criteria were checked using semistructured interviews on sociodemographic factors and medical history. Psychiatric diagnoses were established using the Structured Clinical Interview for Diagnostic and Statistical Manual of Depressive Disorders IV [21]. Depression severity was rated on the IDS-C scale [20] by trained raters. A total of 25 participants were included in the study. To be included in this current analysis, participants had to collect data over a continuous period of at least 130 days. Additionally, subjects were excluded if missing values exceeded 30% during this period. After applying these additional criteria, 22 of 25 participants were included in the present analysis.

App-Based Self-Rating and Measures

According to the goals of the project, an app was developed by the STEADY consortium to support the self-rating process. The study participants performed morning and evening self-ratings using the app. The smartphone app was programmed so that morning logs were available daily from 3 am to 3 pm and evening logs were available daily from 3 pm to 3 am. Patients were asked to fill in the morning logs directly after awakening and to fill in the evening logs right before going to bed.

Two main self-ratings were considered: daily ratings of depression core symptoms and night sleep. The former considered core symptoms addressing interest and pleasure as well as feelings of depression and hopelessness such as those used in the Two-item Health Questionnaire (PHQ-2) [22] to measure depression symptoms. In the evening log, patients rated how often they experienced the symptoms during the day on a visual analogue scale from 0 (“never”) to 10 (“all the time”). A mean of the two items was calculated and used as a measure of depression core symptoms severity, hereinafter referred to as depression core symptoms. For night sleep measurement, participants were asked to fill in a sleep diary that was integrated in the morning log consisting of “go to bed time,” “get up time,” and “time spent sleeping.” Time in bed and total sleep time were then calculated from the sleep diary as the main parameters for analysis.

Procedure

Patients were invited to the study center where they were given a general introduction to the app and the integrated morning and evening logs. Patients were instructed on how and when to fill in the logs. Monthly visits at the study center took place during the data collection period, which involved a data plausibility check (including a check for missing values), along with the opportunity to address patient questions and difficulties. Study assistants monitored the data collection process throughout this period. This involved a continuous check for missing values. When a missing value was noted, it was addressed in a phone call with the patient to obtain feedback on the reasons for missing values as a measure to reduce the occurrence of missing values.

Data Analysis

The vector autoregression (VAR) technique was applied to analyze the collected datasets [23]. The VAR technique was originally developed for research in economic sciences, but has also been used in other fields such as neuroimaging, sociology, and meteorology. A useful advantage of the VAR technique is that it allows investigating the temporal dynamics between several time series without the need of presumptions about the possible associations. In a VAR model, all of the endogenous variables are regressed onto their own time-lagged values and the time-lagged values of all other endogenous variables. An endogenous variable is a variable that can be both a determinant and an outcome. Therefore, a VAR model enables drawing conclusions about the temporal sequence of the effects to derive indications for cause-effect relationships [23].

Several steps are required when conducting a VAR analysis. One important step is to make a decision about the maximum lag. The lag length refers to the maximum number of previous observations (eg, the previous 2 days), which is then used to estimate the current observation. Lag length selection criteria can be used to determine the optimal number of lags.

In this study, we established two models with two variables each. Each model consisted of two endogenous variables: total sleep time and depression core symptoms (model 1), and time in bed and depression core symptoms (model 2). The corresponding regression equations are listed below.

Equations for model 1:

Equations for model 2:

In these models, α_i, βi, γi, and δi are coefficients to be estimated; ε_1 and ε_2 are the error terms; and p is the number of lags considered in the model. A maximal lag length of 7 days was used. The Akaike information criterion value was used to identify the optimal number of lags for the models. To account for trends and to obtain stationarity, the time series were passed through a linear filter using the autoregressive integrated moving average methodology [24]. To test for potential cause-effect associations, Granger causality tests were conducted. This test examines whether the past values of a variable are useful to predict another variable. A variable x is considered to be causally related to a variable y if past values of y and x predict y significantly better than the past values of y alone. The principle behind the Granger test is that a cause cannot come after an effect [15]. Therefore, testing the temporal order of associations provides information on potential causal effects. To determine the impact of each endogenous variable over time, impulse response functions (IRFs) were calculated for all patients with significant Granger causality test results. An impulse response is the temporal reaction of a dynamic system to a change of a variable (eg, change of total sleep time in our analysis). Such a change is called "shock" and is normally defined as 1 SD in the time series; this definition was adopted in the present study as it has been shown to be a reliable definition of change in clinical research [18]. Furthermore, forecast error variance decomposition (FEVD) was performed to estimate the amount of variance in a variable that can be explained by another variable in a defined time period. The R package *vars* was used for statistical analysis [25].

Missing Values

Missing values occurred occasionally in the time series whenever a patient forgot to fill in the morning or evening logs. Since the applied analysis method cannot handle missing data, imputation methods were used to estimate missing values. In a given time series, missing values are estimated on the basis of the respective time series itself to avoid sham correlations between different time series. For this purpose, the Kalman smoothing imputation method was applied using the R package *imputeTS* [26].

Results

Demographics and Clinical Characteristics

Data of 22 participants (15 women, 7 men) were analyzed in this study. On average, data for 173 days (range 143-205) per participant were available. Participants were aged between 21 and 67 years (median 43.5) and had a mean IDS-C baseline score of 27.27 (SD 8.5), corresponding to a moderate severity of depression, and a mean PSQI baseline score of 8.95 (SD 4.62). Nineteen of the 22 participants (86%) were taking antidepressant medication and had already undergone psychotherapy or were in psychotherapy at the time of the study. Table 1 summarizes all demographic and clinical characteristics at the individual patient level.

Table 1. Demographic and clinical characteristics.

Patient	Data collection period (days)	Sex	Age (years)	Employment status	Other psychiatric disorders	PSQI[a]		IDS-C[b]	
						T1[c]	T2[d]	T1	T2
1	201	F	41	Housewife	Anxiety disorder/ agoraphobia	—[e]	11	24	14
2	150	F	50	Full time	Generalized anxiety disorder, May 2017-present	12	13	23	30
3	174	M	54	Incapacity for work due to illness	Generalized anxiety disorder	4	6	39	21
4	179	F	33	Part time	None	5	9	17	18
5	172	F	43	Full time	None	5	4	6	7
6	205	F	60	Partial retirement	None	10	9	34	34
7	199	F	44	Self-employed/ freelance	None	4	6	7	16
8	150	F	45	Full time	Pain disorder, 2010-present	18	14	9	27
9	188	F	38	Full time	None	10	18	18	38
10	156	F	33	Full time	None	7	7	11	10
11	203	M	42	Self-employed/ freelance	None	2	3	19	9
12	170	M	21	Student	None	10	10	11	15
13	154	M	53	Partial retirement	None	22	20	44	44
14	160	F	48	Self-employed/ freelance	None	7	5	15	6
15	143	M	50	Privateer	None	7	7	31	36
16	150	F	35	Sick leave	None	3	7	22	25
17	199	F	32	Full time	None	9	–	23	
18	179	F	53	Part time	Trichotillomania	8	5	15	17
19	178	F	27	Student	Attention deficit and hyperactivity disorder	–	15	20	16
20	173	F	67	Pension/early retirement	None	10	12	29	20
21	178	M	26	Part time	None	7	9	5	14
22	151	M	62	Incapacity for work due to illness	None	16	17	45	39

[a]PSQI: Pittsburg Sleep Quality Index score.

[b]IDS-C: Inventory of Depressive Symptomatology Clinician Rating.

[c]T1: before the study phase.

[d]T2: after the study phase.

[e]—: data not available.

Granger Causality Results

Granger causality tests were performed to investigate the temporal order of associations and to test for potential causal effects. In total, 11 patients showed significant Granger causal associations at a significance level of 5%. For 3 patients, time in bed Granger caused depression core symptoms (positive association), meaning that more time in bed was followed by more severe depression. In 3 patients, this association had a reversed temporal order, meaning that more depression core symptoms were followed by more time in bed. In one patient, depression core symptoms Granger caused time in bed (negative association), meaning that more depression core symptoms were followed by less time in bed. In 3 patients, total sleep time Granger caused depression core symptoms (positive association), meaning that more total sleep time was followed by more depression, whereas this temporal order was reversed in 2 patients (ie, more depression core symptoms were followed by greater total sleep time). In 3 patients, total sleep time Granger caused depression core symptoms (negative association), meaning that more total sleep time was followed by less

depression core symptoms, and this temporal order was reversed for 1 patient. Figure 1 shows the raw time series (time in bed and depression core symptoms) of patient 2 as an example. Table 2 summarizes all of the statistically significant Granger causality test results.

Figure 1. Raw time-series data of patient 2.

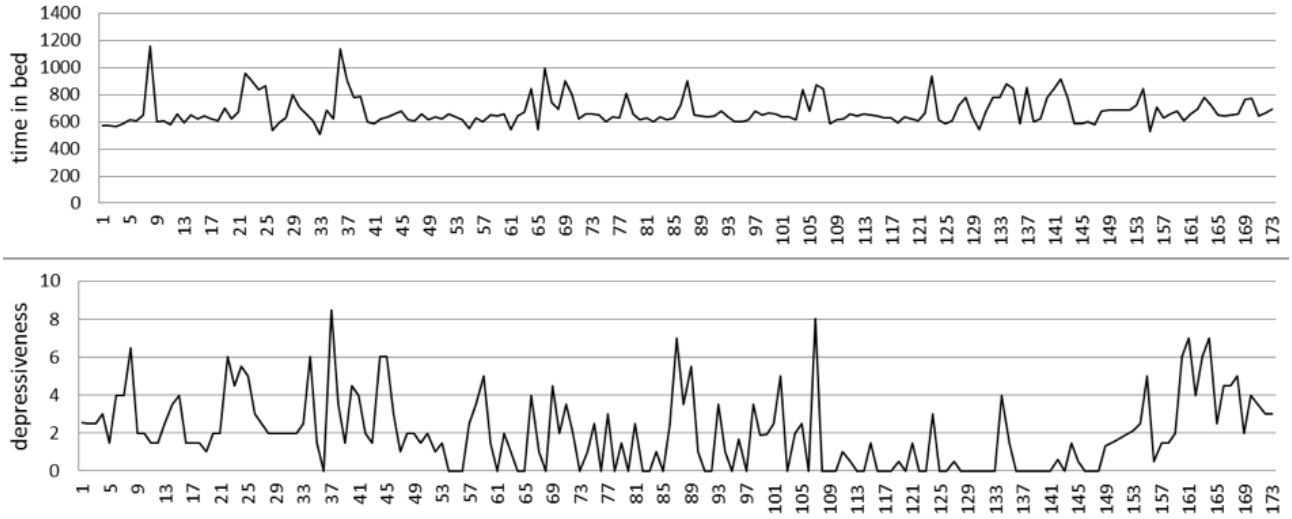


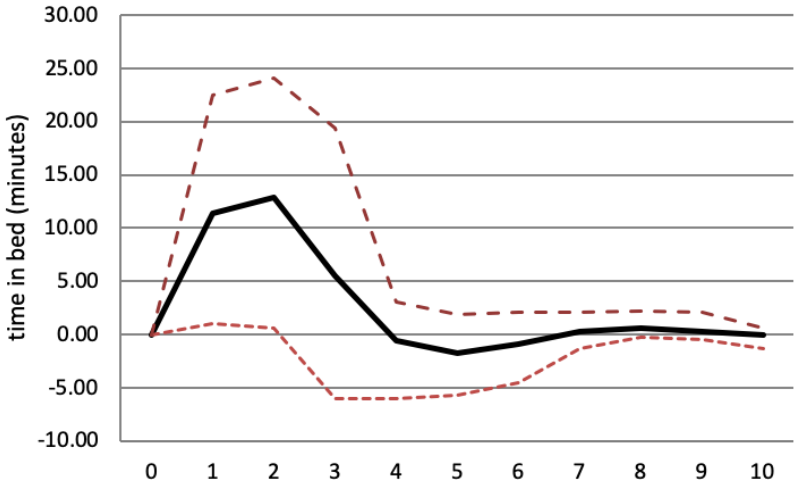
Table 2. Granger causality test results.

Causality test	Positive association (n)	Negative association (n)	Total (n)
Time in bed (TIB)			
TIB Granger causes depression core symptoms	3	0	3
Depression core symptoms Granger causes TIB	3	1	4
Mutual causality	0	0	0
Total sleep time (TST)			
TST Granger causes depression core symptoms	3	3	6
Depression core symptoms Granger causes TST	2	1	3
Mutual causality	0	0	0

Impulse Response Function

IRFs were calculated to analyze the impact of a change in the two sleep variables on depression core symptoms and vice versa over time. Figure 2 provides an example of the IRF for patient 2, showing the impact of a change (1 SD) in depression core symptoms on time in bed during a 10-day period.

Figure 2. Impulse response function for patient 2. Impulse represents depression severity, response is time in bed, time horizon = 10 days; the dashed line indicates the 95% CI.



In this case, a change in depression core symptoms led to a significant increase in time in bed the following 2 days. At day 2, the effect reached its peak with an increase of 12.86 minutes in bed. At day 3, the impact was no longer significant based on a confidence interval exceeding 0. IRF coefficients can be cumulated to calculate the total impact of a change (1 SD) in a variable over time. For patient 2, a change in depression core symptoms of 1 SD led to a total increase in time in bed of 27.64 minutes during a 10-day period. Table 3 summarizes the cumulated IRF results for all patients with significant Granger causality test results.

Table 3. Cumulative results[a] of impulse response analysis for all patients with significant Granger causality test results.

Granger causality	lag[b] 0	lag 1	lag 2	lag 3	lag 4	lag 5	lag 6	lag 7	lag 8	lag 9	lag 10
TIB[c] Granger causes depression											
Patient 3	0.06	0.29	0.51	0.53	0.53	0.53	0.52	0.52	0.52	0.52	0.52
Patient 6	0.02	0.25	0.31	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.31
Patient 20	0.44	0.89	1.00	1.02	1.03	1.03	1.03	1.03	1.03	1.03	1.03
Mean	0.17	0.47	0.60	0.62	0.63	0.62	0.62	0.62	0.62	0.62	0.62
Depression Granger causes TIB											
Patient 10	0.00	7.46	9.17	9.35	9.34	9.33	9.33	9.33	9.33	9.33	9.33
Patient 11	0.00	8.42	9.44	9.72	9.78	9.79	9.79	9.79	9.79	9.79	9.79
Patient 14	0.00	12.74	13.33	13.55	13.57	13.57	13.57	13.57	13.57	13.57	13.57
Patient 19	0.00	–7.16	–20.71	–22.59	–23.64	–23.81	–23.66	–23.63	–23.60	–23.60	–23.60
Mean	0.00	8.95	13.16	13.80	14.08	14.13	14.09	14.08	14.07	14.07	14.07
TST[d] Granger causes depression											
Patient 6	0.06	0.27	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
Patient 7	0.09	–0.16	–0.13	–0.14	–0.14	–0.14	–0.14	–0.14	–0.14	–0.14	–0.14
Patient 14	0.08	–0.11	–0.10	–0.10	–0.10	–0.10	–0.10	–0.10	–0.10	–0.10	–0.10
Patient 19	–0.17	–0.46	–0.48	–0.49	–0.49	–0.49	–0.49	–0.49	–0.49	–0.49	–0.49
Patient 20	–0.24	0.20	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
Patient 21	0.36	0.09	–0.03	–0.10	–0.25	–0.40	–0.40	–0.35	–0.30	–0.27	–0.24
Mean	0.17	0.22	0.22	0.23	0.26	0.28	0.28	0.27	0.26	0.26	0.25
Depression Granger causes TST											
Patient 2	0.00	9.28	20.88	27.44	28.44	27.79	27.23	27.49	28.03	28.44	28.55
Patient 11	0.00	9.11	10.24	10.57	10.63	10.64	10.65	10.65	10.65	10.65	10.65
Patient 22	0.00	–9.46	–9.36	–9.37	–9.37	–9.37	–9.37	–9.37	–9.37	–9.37	–9.37
Mean	0.00	9.28	13.49	15.79	16.15	15.93	15.75	15.84	16.02	16.15	16.19

[a]Values are expressed in minutes (TIB and TST) or units on the depression scale (depression).

[b]Lag numbers correspond to days.

[c]TIB: time in bed.

[d]TST: total sleep time.

Variance Decomposition

Table 4 summarizes the results of the FEVD analysis over a period of 10 days for all patients with significant Granger causality test results. As an example, in patient 20, 6.4% of the variance in depression core symptoms at lag 1 (the day following night sleep) could be explained by time in bed. The following day (lag 2), the percentage increased to 12.4%. At day 3 (lag 3), a peak was reached with 12.7% of the variance explained. The mean FEVD results were calculated for each temporal association across patients with respective significant Granger causal associations. On average, 10% and 6% of the variance in depression core symptoms could be explained by time in bed and total sleep time, respectively, and 3% of variance in time in bed and in total sleep time could be explained by depression core symptoms.

Table 4. Explained variance in depression core symptoms and time in bed at different time lags for all patients with significant Granger causality test results.

Granger causality	lag[a] 1	lag 2	lag 3	lag 4	lag 5	lag 6	lag 8	lag 9	lag 10
TIB[b] Granger causes depression									
Patient 3	0.00	0.07	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Patient 6	0.00	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Patient 20	0.06	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Mean	0.02	0.08	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Depression Granger causes TIB									
Patient 10	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Patient 11	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Patient 14	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Patient 19	0.00	0.01	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Mean	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03
TST[c] Granger causes depression									
Patient 6	0.00	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Patient 7	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Patient 14	0.01	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Patient 19	0.01	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Patient 20	0.02	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Patient 21	0.09	0.13	0.14	0.14	0.15	0.16	0.16	0.16	0.16
Mean	0.02	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Depression Granger causes TST									
Patient 2	0.00	0.02	0.04	0.05	0.05	0.05	0.05	0.05	0.05
Patient 11	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Patient 22	0.00	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Mean	0.00	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03

[a]Lag numbers correspond to days.

[b]TIB: time in bed.

[c]TST: total sleep time.

Discussion

In this study, idiographic analyses were applied on time-series data of self-reported sleep and depression ratings in 22 patients with major depressive disorder. Data were collected on a daily basis with a smartphone app across 143-205 days per subject. Significant Granger causal associations were found in 11 of 22 patients regarding the associations between time in bed or total sleep time and depression core symptoms.

The first aim of this study was to determine the number of patients for whom temporal associations could be found to indicate a causal relationship. In 7 patients, temporal associations between time in bed and depression core symptoms were found to exhibit a causal relationship, and causal associations were found between total sleep time and depression core symptoms in 9 patients.

The second aim of the study was to evaluate whether changes in sleep or changes in depression core symptoms are primarily responsible for causing changes in the other variable, and whether more or less sleep causes a reduction in depression core symptoms. The analysis revealed high heterogeneity in these associations among patients. One source of heterogeneity was related to the temporal order of the associations, and hence the direction of a possible causal relationship (Granger causality). Concerning time in bed, the direction of the association was quite equally distributed: time in bed Granger caused depression core symptoms in 3 patients, whereas depression core symptoms Granger caused time in bed in 4 patients. With regard to total sleep time, the association was more homogenous toward total sleep time being the cause rather than the effect: total sleep time Granger caused depression core symptoms in 6 patients, whereas the temporal order was reversed in 3 patients.

A different source of heterogeneity concerns the nature of the relationship (eg, if more sleep or depression core symptoms

causes a reduction or an increase in the other variable). The association between time in bed and depression core symptoms was positive in 6 patients, meaning that longer time in bed/more depression core symptoms led to an increase in depression core symptoms/time in bed, whereas this association was negative in 1 patient, meaning that more depression core symptoms led to a decrease in time in bed. A more homogenous pattern was found with regard to the association between total sleep time and depression core symptoms: The association was positive in 5 patients, meaning that longer total sleep time/more depression core symptoms led to more depression core symptoms/longer total sleep time, and the association was negative in 4 patients, meaning that longer total sleep time/depression core symptoms led to a decrease in depression core symptoms/total sleep time.

The third aim was to analyze the impact and temporal dynamics of the effects identified. Variance decomposition revealed that, on average, the largest proportion of explained variance was in the association between time in bed and depression core symptoms, with an average 10% of depression core symptoms being explained by time in bed. The second largest proportion concerned the association between total sleep time and depression core symptoms with 6% of the variance in depression core symptoms being explained by total sleep time. Explained variance was small in associations with a reversed temporal order: 3% of the variance in total sleep time and 3% of the variance in time in bed could be explained by depression core symptoms. Impulse response analysis revealed that changes of 1 SD in a variable (eg, total sleep time) had the greatest impact on the other variable (eg, depression core symptoms) in the following 2 to 4 days.

An advantage of our analysis is that it allows the possibility to make statistical claims on an individual level owing to the use of time series. This would be particularly useful for a clinician since it would be highly relevant to know whether a temporal association between sleep and depression exists in a certain patient and whether it is more likely that sleep changes cause changes in depression or vice versa. If patients are willing to share such data with a clinician [27], it could be used for treatment decisions. For example, if the data suggest that a longer time in bed worsens depressive symptoms, therapeutic sleep restriction or long-term mild sleep restriction might be a promising intervention. This knowledge could also be used by the patients themselves for self-management activities.

Acknowledgments

This study was supported by a grant of the Federal Ministry of Education and Research of Germany (project number 13GW0162B/13GW0162A). The authors assume all responsibility for the content of this publication.

Conflicts of Interest

NL is the founder and shareholder of the company mementor GmbH that developed an online sleep training app to treat insomnia. The remaining authors declare no conflicts of interest.

References

1. Hertenstein E, Feige B, Gmeiner T, Kienzler C, Spiegelhalder K, Johann A, et al. Insomnia as a predictor of mental disorders: A systematic review and meta-analysis. Sleep Med Rev 2019 Feb;43:96-105. [doi: 10.1016/j.smrv.2018.10.006] [Medline: 30537570]
2. Fernandez-Mendoza J, Shea S, Vgontzas AN, Calhoun SL, Liao D, Bixler EO. Insomnia and incident depression: role of objective sleep duration and natural history. J Sleep Res 2015 Aug;24(4):390-398. [doi: 10.1111/jsr.12285] [Medline: 25728794]
3. Lee E, Cho HJ, Olmstead R, Levin MJ, Oxman MN, Irwin MR. Persistent sleep disturbance: a risk factor for recurrent depression in community-dwelling older adults. Sleep 2013 Nov 01;36(11):1685-1691 [FREE Full text] [doi: 10.5665/sleep.3128] [Medline: 24179302]
4. Maglione JE, Ancoli-Israel S, Peters KW, Paudel ML, Yaffe K, Ensrud KE, Study of Osteoporotic Fractures Research Group. Subjective and objective sleep disturbance and longitudinal risk of depression in a cohort of older women. Sleep 2014 Jul 01;37(7):1179-1187 [FREE Full text] [doi: 10.5665/sleep.3834] [Medline: 25061246]
5. Bao Y, Han Y, Ma J, Wang R, Shi L, Wang T, et al. Cooccurrence and bidirectional prediction of sleep disturbances and depression in older adults: Meta-analysis and systematic review. Neurosci Biobehav Rev 2017 Apr;75:257-273. [doi: 10.1016/j.neubiorev.2017.01.032] [Medline: 28179129]
6. Alvaro PK, Roberts RM, Harris JK. A Systematic Review Assessing Bidirectionality between Sleep Disturbances, Anxiety, and Depression. Sleep 2013 Jul 01;36(7):1059-1068 [FREE Full text] [doi: 10.5665/sleep.2810] [Medline: 23814343]
7. Lorenz N, Heim E, Roetger A, Birrer E, Maercker A. Randomized Controlled Trial to Test the Efficacy of an Unguided Online Intervention with Automated Feedback for the Treatment of Insomnia. Behav Cogn Psychother 2019 May;47(3):287-302. [doi: 10.1017/S1352465818000486] [Medline: 30185239]
8. Christensen H, Batterham PJ, Gosling JA, Ritterband LM, Griffiths KM, Thorndike FP, et al. Effectiveness of an online insomnia program (SHUTi) for prevention of depressive episodes (the GoodNight Study): a randomised controlled trial. Lancet Psychiatry 2016 Apr;3(4):333-341. [doi: 10.1016/s2215-0366(15)00536-2]

9. Blom K, Jernelöv S, Kraepelien M, Bergdahl MO, Jungmarker K, Ankartjärn L, et al. Internet treatment addressing either insomnia or depression, for patients with both diagnoses: a randomized trial. Sleep 2015 Feb 01;38(2):267-277 [FREE Full text] [doi: 10.5665/sleep.4412] [Medline: 25337948]
10. Hegerl U, Hensch T. The vigilance regulation model of affective disorders and ADHD. Neurosci Biobehav Rev 2014 Jul;44:45-57. [doi: 10.1016/j.neubiorev.2012.10.008] [Medline: 23092655]
11. Reynold AM, Bowles ER, Saxena A, Fayad R, Youngstedt SD. Negative Effects of Time in Bed Extension: A Pilot Study. J Sleep Med Disord 2014 Aug 28;1(1):1002 [FREE Full text] [Medline: 25374963]
12. Giedke H, Wormstall H, Haffner H. Therapeutic sleep deprivation in depressives, restricted to the two nocturnal hours between 3:00 and 5:00. Prog Neuro-Psychoph 1990 Jan;14(1):37-47. [doi: 10.1016/0278-5846(90)90062-l]
13. Bauer M, Grof P, Rasgon N, Bschor T, Glenn T, Whybrow PC. Temporal relation between sleep and mood in patients with bipolar disorder. Bipolar Disord 2006 Apr;8(2):160-167. [doi: 10.1111/j.1399-5618.2006.00294.x] [Medline: 16542186]
14. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003;54:579-616. [doi: 10.1146/annurev.psych.54.101601.145030] [Medline: 12499517]
15. Granger CWJ. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. Econometrica 1969 Aug;37(3):424. [doi: 10.2307/1912791]
16. Molenaar PC, Campbell CG. The New Person-Specific Paradigm in Psychology. Curr Dir Psychol Sci 2009 Apr;18(2):112-117. [doi: 10.1111/j.1467-8721.2009.01619.x]
17. van der Krieke L, Emerencia AC, Bos EH, Rosmalen JG, Riese H, Aiello M, et al. Ecological Momentary Assessments and Automated Time Series Analysis to Promote Tailored Health Care: A Proof-of-Principle Study. JMIR Res Protoc 2015 Aug 07;4(3):e100 [FREE Full text] [doi: 10.2196/resprot.4000] [Medline: 26254160]
18. Rosmalen JGM, Wenting AMG, Roest AM, de JP, Bos EH. Revealing causal heterogeneity using time series analysis of ambulatory assessments: application to the association between depression and physical activity after myocardial infarction. Psychosom Med 2012 May;74(4):377-386. [doi: 10.1097/PSY.0b013e3182545d47] [Medline: 22582335]
19. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
20. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychol Med 2009 Jul 09;26(3):477-486. [doi: 10.1017/S0033291700035558] [Medline: 30886898]
21. First MB, Gibbon M. The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In: Comprehensive Handbook Of Psychological Assessment, Volume 2: Personality Assessment. Hoboken, NJ: John Wiley & Sons; 2004:134-143.
22. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res 2005 Feb;58(2):163-171. [doi: 10.1016/j.jpsychores.2004.09.006] [Medline: 15820844]
23. Brandt PT, Williams JT. Multiple time series models. In: Quantitative Applications In The Social Sciences. Thousand Oaks, CA: Sage Publications Inc; 2020.
24. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. Time Series Analysis: Forecasting And Control (Wiley Series In Probability And Statistics). Hoboken, NJ: John Wiley & Sons; 2020.
25. Pfaff B, Stigler M. Cran Network. VAR Modelling URL: https://cran.r-project.org/web/packages/vars/vars.pdf [accessed 2019-11-15]
26. Moritz S, Bartz-Beielstein T. imputeTS: Time Series Missing Value Imputation in R. R Journal 2017;9(1):207. [doi: 10.32614/rj-2017-009]
27. Hartmann R, Sander C, Lorenz N, Böttger D, Hegerl U. Utilization of Patient-Generated Data Collected Through Mobile Devices: Insights From a Survey on Attitudes Toward Mobile Self-Monitoring and Self-Management Apps for Depression. JMIR Ment Health 2019 Apr 03;6(4):e11671 [FREE Full text] [doi: 10.2196/11671] [Medline: 30942693]

Abbreviations

FEVD: forecast error variance decomposition
IDS-C: Inventory of Depressive Symptomatology Clinician
IRF: impulse response function
PHQ-2: Two-item Health Questionnaire
PSQI: Pittsburg Sleep Quality Index
STEADY: sensor-based system for therapy support and management of depression
VAR: vector autoregression

Edited by G Eysenbach; submitted 15.11.19; peer-reviewed by E Lee, M Spiliopoulou, M Lang; comments to author 08.12.19; revised version received 04.01.20; accepted 25.03.20; published 23.04.20.

Please cite as:
Lorenz N, Sander C, Ivanova G, Hegerl U
Temporal Associations of Daily Changes in Sleep and Depression Core Symptoms in Patients Suffering From Major Depressive Disorder: Idiographic Time-Series Analysis
JMIR Ment Health 2020;7(4):e17071
URL: http://mental.jmir.org/2020/4/e17071/
doi:10.2196/17071
PMID:32324147

Original Paper

A Web-Based Adaptation of the Quality of Life in Bipolar Disorder Questionnaire: Psychometric Evaluation Study

Emma Morton[1,2], BSc (Hons), PhD; Sharon HJ Hou[3], MA, BA; Oonagh Fogarty[4], BSc; Greg Murray[2], BSc, BA (Hons), MPsych, PhD; Steven Barnes[5], PhD; Colin Depp[6], PhD; CREST.BD[1]; Erin Michalak[1], PhD

[1]Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

[2]Faculty of Health, Arts and Design, Swinburne University, Hawthorn, Australia

[3]University of Guelph, Guelph, ON, Canada

[4]Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

[5]Department of Psychology, University of British Columbia, Vancouver, BC, Canada

[6]Department of Psychiatry, University of California, San Diego, CA, United States

Corresponding Author:
Emma Morton, BSc (Hons), PhD
Department of Psychiatry
University of British Columbia
420-5950 University Blvd
Vancouver, BC, V6T 1Z3
Canada
Phone: 1 604 827 3393
Email: emma.morton@ubc.ca

Abstract

Background: Quality of life (QoL) is considered a key treatment outcome in bipolar disorder (BD) across research, clinical, and self-management contexts. Web-based assessment of patient-reported outcomes offer numerous pragmatic benefits but require validation to ensure measurement equivalency. A web-based version of the Quality of Life in Bipolar Disorder (QoL.BD) questionnaire was developed (QoL Tool).

Objective: This study aimed to evaluate the psychometric properties of a web-based QoL self-report questionnaire for BD (QoL Tool). Key aims were to (1) characterize the QoL of the sample using the QoL Tool, (2) evaluate the internal consistency of the web-based measure, and (3) determine whether the factor structure of the original version of the QoL.BD instrument was replicated in the web-based instrument.

Methods: Community-based participatory research methods were used to inform the development of a web-based adaptation of the QoL.BD instrument. Individuals with BD who registered for an account with the QoL Tool were able to opt in to sharing their data for research purposes. The distribution of scores and internal consistency estimates, as indicated by Cronbach alpha, were inspected. An exploratory factor analysis using maximum likelihood and oblique rotation was conducted. Inspection of the scree plot, eigenvalues, and minimum average partial correlation were used to determine the optimal factor structure to extract.

Results: A total of 498 people with BD (349/498, 70.1% female; mean age 39.64, SD 12.54 years; 181/498, 36.3% BD type I; 195/498, 39.2% BD type II) consented to sharing their QoL Tool data for the present study. Mean scores across the 14 QoL Tool domains were, in general, significantly lower than that of the original QoL.BD validation sample. Reliability estimates for QoL Tool domains were comparable with that observed for the QoL.BD instrument (Cronbach alpha=.70-.93). Exploratory factor analysis supported the extraction of an 11-factor model, with item loadings consistent with the factor structure suggested by the original study. Findings for the sleep and physical domains differed from the original study, with this analysis suggesting one shared latent construct.

Conclusions: The psychometric properties of the web-based QoL Tool are largely concordant with the original pen-and-paper QoL.BD, although some minor differences in the structure of the sleep and physical domains were observed. Despite this small variation from the factor structure identified in the QoL.BD instrument, the latent factor structure of the QoL Tool largely reproduced the original findings and theoretical structure of QoL areas relevant to people with BD. These findings underscore the research and clinical utility of this instrument, but further comparison of the psychometric properties of the QoL Tool relative

to the QoL.BD instrument is warranted. Future adaptations of the QoL Tool, including the production of an app-based version of the QoL Tool, are also discussed.

(JMIR Ment Health 2020;7(4):e17497) doi:10.2196/17497

KEYWORDS

bipolar disorder; survey methodology; patient reported outcomes; psychometrics; questionnaire design; quality of life; validation studies

Introduction

Background

Applications of quality of life (QoL) assessment instruments in bipolar disorder (BD) research have grown substantially [1,2] since the introduction of the concept in psychiatric research more generally in the 1980s [3]. Broadly speaking, QoL instruments holistically assess an individual's satisfaction and functioning across a range of life domains and are, therefore, increasingly used to evaluate BD treatment outcomes beyond symptomatic response. Assessment of QoL may be particularly important in the context of BD, given the chronic course and significant impacts across diverse life domains associated with this mood disorder [1]. Indeed, there is some evidence to indicate that both patients with BD and health care providers view QoL as the most important outcome in the treatment of the condition [4].

In the study of BD, QoL has been primarily measured with universal or generic instruments (most commonly, the 36-Item Short Form Health Survey and Quality of Life Enjoyment and Satisfaction Questionnaire [1]). Although generic measures assess areas of life, which may be considered fundamentally important [5], patient groups may have unique priorities that are best assessed with disorder-specific instruments [6]. To address this gap, the first condition-specific QoL instrument for BD was developed: the Quality of Life in Bipolar Disorder (QoL.BD) [7]. Informed by consultation with people with lived experience of BD, their family members, and field experts, the resulting scale assesses cardinal life areas directly impacted by BD symptoms (mood, sleep, physical health, and cognition), pragmatic and functional outcomes (home, work, education, leisure, and finances), and more psychosocially orientated constructs (relationships, self-esteem, spirituality, identity, and independence). A decade since its development, the QoL.BD instrument has seen international adoption: it has undergone formal adaptation and validation in Iranian, Chinese, and Chilean populations [8-10] and has been translated into over 20 languages [11]. It has also seen application in diverse research contexts, including clinical trials of psychotherapy [12-16] and pharmacological interventions [17,18]. Materials have also been developed to support the use of the QoL.BD instrument by health care practitioners in a clinical context (eg, case formulation [11]) and by individuals with BD themselves in their self-management practices [19].

Although the uptake of the QoL.BD instrument has been encouraging, its research and clinical utility may be enhanced with a web-based delivery format. Relative to traditional pen-and-paper instruments, web-based administration formats reduce administrative burden (through, for instance, automatic scoring and practical data storage), data entry and coding errors, and item nonresponse [20]. Web-based questionnaires may also enhance the accessibility of instruments for both researchers and patients: they are cost-effective [21], instantaneously available to potential users with an internet connection (regardless of location), and navigation is user-friendly, with the ability to skip or eliminate irrelevant questions from the view [22]. Respondents may also prefer web-based administration formats [23], and for questionnaires that assess sensitive topics (such as factors related to mental health), web-based questionnaires may potentially reduce social desirability effects [24]. For ongoing self-monitoring purposes, web-based instruments are advantaged by their ability to provide immediate feedback to respondents, reduce the burden of tracking large volumes of data, and potentially lessen experiences of stigma by decreasing the visibility of symptom monitoring. Given the numerous pragmatic benefits and enhanced user-friendliness of web-based self-report questionnaires, adaptation of the pen-and-paper version of the QoL.BD instrument to a web-based interface was undertaken to support utilization of this instrument across research, clinical, and self-management contexts.

However, simple migration of pen-and-paper scales to web-based formats does not guarantee preservation of a scale's psychometric properties. A number of factors can impact the way a scale performs when adapted for web-based administration, including modifications to layout, instructions, or changes in item wording and response options [25,26]. The Professional Society for Health Economics and Outcomes Research (ISPOR) guidelines suggest that evidence needed to support measurement equivalence between pen-and-paper and electronic adaptations varies depending on the extent of modifications, from minor (eg, simply displaying a scale text on screen) to more substantive (ranging from moderate alterations such as splitting the presentation of items over several screens, up to large scale changes to items, presentation or response format). Supporting evidence can include usability testing, appraisal of interformat reliabilities, comparable means and standard deviations, and preservation of scale reliability and factor structures across formats. Although the majority of Web-adaptation studies have reported interformat reliabilities, informing confidence about the consistency of measurement of self-reported mental health data across formats [27,28], fewer studies have made a comment on whether the original factor structure is replicated in web-based questionnaire formats. As such, exploration of the psychometric properties, particularly factor structure, is needed to support the use of any web-based adaptation of the QoL.BD instrument.

Objective

The overarching aim of this study was to compare the psychometric performance of the web-based QoL Tool with the original pen-and-paper version of the QoL.BD scale. To do this, we aimed to (1) describe the means and standard deviations of QoL Tool responses, (2) evaluate the internal consistency of the web-based measure, and (3) determine whether the factor structure of the original QoL.BD could be replicated in the web-based adaptation of this instrument.

Methods

Overview

The project was conducted by the Collaborative RESearch Team to study psychosocial issues in Bipolar Disorder (CREST.BD [19]), a Canadian-based network dedicated to collaborative research and knowledge translation (KT) in BD. CREST.BD specializes in community-based participatory research (CBPR), where researchers and knowledge users work collaboratively [29]. Informed by a decade of research and integrated KT, CREST.BD has developed a specific model of CBPR for BD [30]. Funding from the Canadian Institutes of Health Research was granted to extend on prior work to design and validate a pen-and-paper QoL questionnaire for BD (described below). This psychometric evaluation follows the development of a web-based adaptation of the QoL.BD instrument using CBPR methods.

Design and Validation of the Pen-and-Paper Quality of Life in Bipolar Disorder Questionnaire

The development and validation of the QoL.BD instrument is described in detail elsewhere [7]. In brief, candidate items were generated through (1) qualitative interviews with people with BD, their family members, and field experts and (2) a literature review of existing research on QoL in BD. Following item reduction, preliminary psychometric analyses, and further consultation with field and lived experience experts, a final subset of 56 items was retained. Items are organized into 14 4-item domains: 12 core (physical, sleep, mood, cognitive, leisure, social, spirituality, finances, household, self-esteem, independence, and identity) and 2 optional (work and study, which respondents are directed to complete if they are currently employed or in school). A 12-item brief version was also developed.

The questionnaire items are presented on a standard 5-point Likert response scale (strongly disagree–strongly agree). Items are all positively worded (ie, describing the presence of a desirably quality) for two reasons: (1) a positive question frame is consistent with the strengths-based approach to QoL adopted as a result of CBPR consultation and (2) reverse worded items can reduce the reliability and validity of a scale [31]. Domains are scored by summing the responses, for a potential score range of 4 to 16. Calculation of an overall QoL score is possible by summing responses to the 12 core domains. Initial field testing of the QoL.BD instrument indicated that both the full and brief version of the instrument represent a feasible, reliable, and valid BD-specific QoL measure with solid internal validity and appropriate test-retest reliability. Factor analysis affirmed that the 12 basic scales were represented in the latent structure of the instrument.

Development of the Web-Based Adaptation: The Quality of Life Tool

A synergistic combination of CBPR and the principles of user-centered design [32] were applied to develop the QoL Tool. The primary goal of the development process for the QoL Tool was to produce a web-based version of the QoL.BD instrument that was faithful as possible to the original measurement principles of the QoL.BD instrument but also adapted and expanded to enhance user experience and functionality. A priori, we established which features of the QoL.BD instrument were immutable, specifically, preservation of precise wording of the scale's 56 items (with one exception, described below), the ordering of the items, and the 5-point Likert response scale (1=strongly disagree; 5=strongly agree) and item ordering. The approach to scoring the domains and the range of potential scores are consistent with the QoL.BD instrument.

Beyond these parameters, however, it was expected that the web-based version of the scale would differ in some aspects from its pen-and-paper counterpart. One adaptation was made in the delivery format of the web-based version on the basis of user feedback, the name of each domain was made visible to the user (see Figure 1). Furthermore, as the inclusion of graphical feedback of results has been described as a highly prioritized feature for self-management apps for people with BD [32,33], we determined *a priori* that the addition of a results display feature would be essential (see Figure 2). All other adaptations were identified via user-centered design processes. One minor change in wording was made to a sleep domain item ("woken up" was changed to "awoken"). All other items in the QoL Tool were precisely as worded in the QoL.BD instrument. Substantial changes were made in the web-based version in terms of features and functionality. For example, registered users of the QoL Tool are provided with the option of an interactive results feature that demonstrates their QoL scores over time (Figure 3) and the option to email their results to a health care provider.

Figure 1. QoL Tool questionnaire screen and response options.

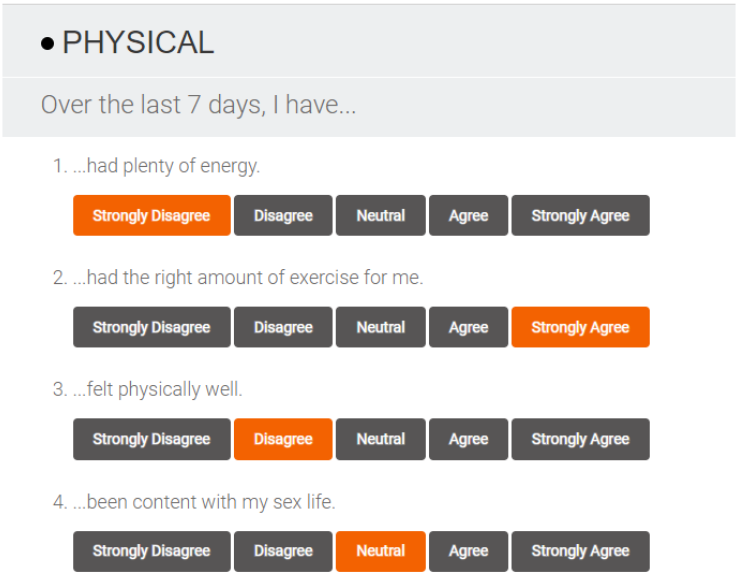


Figure 2. QoL Tool graphical display of results.

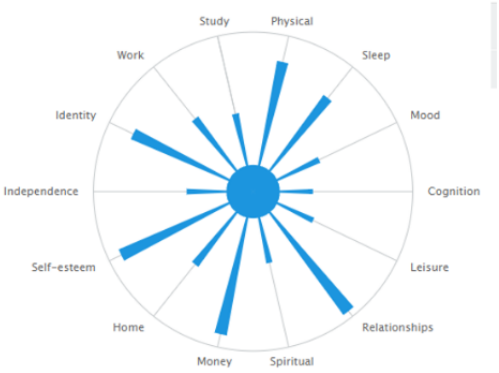
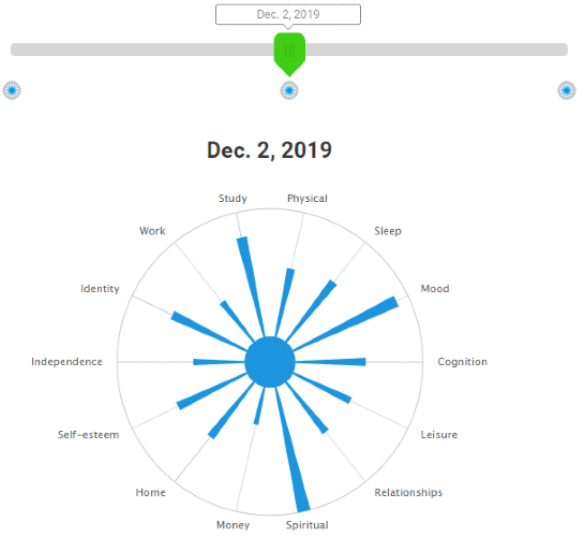


Figure 3. Users are able to drag a slider to compare their QoL Tool results over time.



Recruitment

The final version of the QoL Tool was formally launched on World Bipolar Day (March 30, 2015). The QoL Tool was promoted primarily through social media channels (eg, CREST.BD Facebook page, Twitter, and website) but was also highlighted as part of a series of knowledge translation events (May-June 2015) focused on sharing knowledge on self-management strategies for BD [34]. Informed consent for data collection for research purposes was provided at the point of registration in the QoL Tool system but was not required for the use of the web-based interface. Inclusion criteria were (1) older than 19 years; (2) have a self-reported diagnosis of BD I, II, or not otherwise specified (NOS); and (3) able to communicate in English. No compensation was offered for participants' completion of the QoL Tool. Ethics approval was obtained from the University of British Columbia's Behavior Research Ethics Board.

Data Collection

Data collection for this study occurred between March 30, 2015, and October 17, 2018. Demographic details and responses to the QoL Tool were saved in a secure database hosted at the University of British Columbia. For the purposes of this study, only baseline responses to the QoL Tool were analyzed to avoid contaminating the extracted factor structure with potential learning effects [35]. Data from the baseline time point of the original QoL.BD validation study (n=224; sample described [7]) were included where relevant for comparison purposes.

Statistical Analysis

Internal Consistency

Internal consistency was evaluated using Cronbach alpha for each of the 14 domains. Scales were deemed to be adequately reliable if Cronbach alpha exceeded .7 [36].

Exploratory Factor Analysis

Analyses were carried out on the 12 basic domains of the web-based QoL.BD ie, all except work and study). Factorability was confirmed by using the Kaiser-Meyer-Oklin measure of sampling adequacy and Bartlett test of sphericity. No corrections for missing data were required, as the design of the QoL Tool does not permit users to submit their data unless responses have been provided for all questions. An exploratory factor analysis (EFA) using maximum likelihood extraction was conducted in SPSS 27 (IBM, Armonk, New York). Given that the variables were assumed to be correlated, oblique (oblimin) rotation was applied. Multiple criteria were reviewed to determine the optimal number of factors retained. First, visual inspection of the point of inflexion displayed by the scree plot was conducted [37]. Second, Kaiser criterion was used to determine the number of factors with eigenvalues with a value greater than 1 [38]. Third, the minimum average partial correlation (MAP) test was applied using SPSS [39] to identify the number of components that produces the minimum mean squared partial correlation [40]. Finally, the percentage of variance explained by each factor was considered: amount of variance explained (with 5% a generally accepted cutoff) may be used as a decision rule [41,42].

Interpretability of the extracted factors was evaluated by confirming primary factor loadings exceeded 0.4 and that cross-loadings were less than 0.3 [42]. Finally, the item content of each factor was evaluated to confirm whether the domains proposed by the original validation of the QoL.BD instrument were represented.

Results

Participants

A final sample of 498 participants (349/498, 70.1% female; 128/498, 25.7% male; 6/498, 1.2% transgender or nonbinary) with a mean age of 39.64 years (SD 12.54) were included in the analysis. In total, 36.3% (181/498) of the participants reported having a diagnosis of BD type I (BD-I), 39.2% (195/498) reported having a diagnosis of BD type II (BD-II), and 2.2% (11/498) self-identified as having a diagnosis of BD NOS. Individuals with other unspecified BD (50/498, 10.5%), unclear or pending diagnoses (31/498, 6.5%), or rapid-cycling BD (7/498, 1.5%) comprised the remainder. The majority of participants were located in North America (22/498, 44.2% Canadian; 138/498, 27.7% American), with the remainder comprising international respondents (most commonly from Australia with 28/498, 5.6% or Germany with 16/498, 3.2%). Characteristics of this sample (QoL Tool respondents) and the original pen-and-paper QoL.BD validation sample can be found in Table 1. The two samples did not significantly differ with respect to gender composition, X^2_1=1.6 (N=690), P=.21; nor age, t_{720}=1.31, P=.19. The two samples did differ with respect to diagnosis, X^2_1=53.7 (N=590), P<.001; with the QoL Tool sample containing a greater proportion of individuals with BD-II than the pen-and-paper sample.

Table 1. Sample characteristics of the quality of life (QoL) Tool (n=498) and Quality of Life in Bipolar Disorder (n=224) validation sample.

Sample characteristics	QoL Tool validation sample (n=498)	Original Quality of Life in Bipolar Disorder validation sample (n=224)
Gender, n (%)		
Male	128 (25.7)	67 (29.9)
Female	349 (70.1)	146 (65.2)
Trans or nonbinary	6 (1.2)	N/A[a]
Age (years), mean (SD)	39.64 (12.54)	41.00 (13.67)
Diagnosis, n (%)		
BD-I[b]	181 (36.3)	169 (75.4)
BD-II[c]	195 (39.2)	45 (20.1)
Work/employment status, n (%)		
Currently engaged in paid or volunteer employment	317 (63.7)	121 (54.0)
In education	154 (30.9)	52 (23.2)

[a]N/A: not applicable.

[b]BD-I: bipolar disorder type I.

[c]BD-II: bipolar disorder type II.

Distributions by Domain

Mean, standard deviation, and skew of the 14 domains of the QoL Tool are presented in Table 2, along with comparison data from the first time point of the original QoL.BD validation study. The optional work and study sections were completed by 63.7% (317/498) and 30.9% (154/498) of the web-based sample and 54% (121/225) and 23.2% (52/224) of the pen-and-paper sample, respectively. The distribution of QoL Tool domain scores was approximately normal, with all skew values well under the recommended absolute value of 2 [43] and no evidence of floor or ceiling effects. Mean scores for the web-based sample were significantly lower across all domains except finance, relative to the pen-and-paper sample.

Table 2. Descriptive statistics for the 12 basic and two optional domains of the quality of life (QoL) Tool.

Domain	QoL Tool		Quality of life in Bipolar Disorder		*t* test value (*df*)	*P* value
	Mean (SD)	Skew	Mean (SD)	Skew		
Physical	9.61 (3.70)	0.37	11.77 (4.01)	0.01	7.27 (740)	<.001
Sleep	10.05 (3.64)	0.25	11.42 (4.12)	0.04	5.79 (740)	<.001
Mood	10.74 (3.89)	0.05	12.98 (3.98)	−0.27	7.31 (740)	<.001
Cognitive	10.91 (3.68)	0.06	12.78 (4.22)	−0.26	6.19 (740)	<.001
Leisure	11.79 (3.99)	−0.17	13.11 (4.16)	−0.24	4.17 (740)	<.001
Social	12.83 (4.12)	−0.36	14.31 (4.00)	−0.67	4.64 (740)	<.001
Spirituality	11.49 (3.97)	−0.05	13.05 (4.07)	−0.39	4.99 (740)	<.001
Finances	12.42 (4.89)	−0.13	12.67 (4.64)	−0.28	0.67 (740)	.51
Household	10.91 (4.27)	0.06	12.96 (4.11)	−0.32	6.22 (740)	<.001
Self-esteem	12.90 (3.64)	−0.35	13.82 (3.70)	−0.44	3.22 (740)	.001
Independence	14.51 (3.35)	−0.70	15.78 (3.23)	−0.84	4.91 (740)	<.001
Identity	11.15 (4.10)	0.07	13.68 (4.14)	−0.25	7.87 (740)	<.001
Work (optional)	12.83 (4.25)	−0.32	15.20 (3.52)	−0.96	5.46 (436)	<.001
Study (optional)	11.56 (4.26)	−0.02	13.92 (4.96)	−0.48	3.31 (204)	.01

Internal Consistency

Acceptable to excellent reliability estimates were observed for all 14 QoL Tool domains (see Table 3 for Cronbach alpha values for the QoL Tool and comparison data from the first time point of the original QoL.BD validation study). Reliability estimates for the QoL Tool were comparable with those reported for the QoL.BD instrument across all domains.

Table 3. Internal consistency estimates for the quality of life (QoL) Tool and Quality of Life in Bipolar Disorder.

Domain	QoL Tool		Quality of Life in Bipolar Disorder	
	Cronbach alpha	n	Cronbach alpha	n
Physical	.70	498	.79	218
Sleep	.77	498	.83	208
Mood	.88	498	.90	220
Cognitive	.83	498	.91	219
Leisure	.89	498	.91	193
Social	.86	498	.88	220
Spirituality	.91	498	.93	214
Finances	.89	498	.88	214
Household	.93	498	.91	217
Self-esteem	.84	498	.88	221
Independence	.76	498	.81	217
Identity	.86	498	.90	220
Work	.90	317	.89	121
Study	.89	154	.95	52

Exploratory Factor Analysis

The suitability of the data for factor analysis was confirmed, with a "very good" sample size [44] and appropriate factorability. The Bartlett test of sphericity was significant, c^2_{1128}=15,621.64; *P*<.001, and the Kaiser-Meyer-Olkin measure of sampling adequacy was .93, exceeding the recommended minimum value of .60 [35].

Visual inspection of the scree plot using the William guidelines [37] suggested an 11-factor structure. The Kaiser criterion suggested an 11-factor structure, which accounted for 70.48% of the variance. The MAP test identified that the extraction of 12 components was required to produce the minimum mean squared partial correlation. Weighting these findings together, both an 11-factor structure and a 12-factor structure were considered. Owing to the fact that retention of the 12th factor explained less than 2% of the additional variance (below the 5% cutoff used for factor retention [41]), a final 11-item factor structure was retained.

The interpretability of the 11 extracted factors was supported, with the majority of factors (n=9) having at least four items with factor loadings above 0.4. Only one item ("I have felt emotionally balanced") was observed to have a cross-loading above 0.3 (primary loading mood, secondary loading cognition). The pattern matrix with oblique rotation and factor loadings above 0.3 is shown in Multimedia Appendix 1. The item content of the extracted factors largely aligned with the factor structure suggested in the original validation study, as well as the conceptual labeling of the domains, with the exception of items belonging to the sleep and physical domains. The rotated factor structure suggested that a single latent factor best explained the variance of items belonging to these domains. Furthermore, two items from the original physical domain ("I have had the right amount of exercise for me" and "I have been content with my sex life") did not demonstrate significant loadings on any extracted factor.

Discussion

Principal Findings

This study describes the psychometric properties of a web-based QoL questionnaire for individuals with BD (the QoL Tool) adapted from a pen-and-paper measure (QoL.BD [7]). Distributions of the core 12 QoL Tool domains were comparable with those found using the QoL.BD instrument in the original validation sample (although QoL Tool respondents reported lower QoL across the majority of domains), standards for appropriate internal consistency were met, and EFA suggested a factor structure that is adequately concordant with the full pen-and-paper version.

EFA of the core item set of the QoL Tool suggested an 11-factor latent structure, accounting for a similar proportion of variance (70.48%) to the factor structure identified in the original QoL.BD validation study (12-factor structure accounting for 71% of variance). Furthermore, the same items (with the exception of certain sleep and physical items) were observed to load on the domains identified by the original validation study and conceptual structure of the QoL.BD instrument, suggesting that the same constructs are being measured by the web-based and paper-based versions of this instrument [35].

A range of data suggests that, not surprisingly, directly copying a paper questionnaire into a web page results in negligible change to its psychometric properties. A systematic review conducted by Alfonsson et al [27] on the adaptation of 40 symptom scales into digital format indicated that most web-based instruments appear reliable across administration formats. More specifically, van Ballegooijen et al [28] conducted a focused review of the psychometric data of digitized paper

questionnaires measuring symptoms of mood and anxiety disorders, demonstrating adequate psychometric properties of the tools in their web-based formats. Despite a growing body of evidence comparing the psychometric properties of web-based adaptations of questionnaires, three notable limitations of this body of work exist. First, studies have typically examined the interformat reliability (Web-based vs pen-and-paper) of psychosocial instruments. Second, those which have validated web-based mental health measures have typically used general population samples rather than testing the instrument in a clinical population. Third, few studies have reported the type and extent of modifications made in adapting pen-and-paper questionnaires to a web-based format [27], limiting the ability to make inferences about the effect of delivery mode on degree of similarity. Consequently, this study contributes some initial evidence (couched in the limitations discussed below) supporting that factor structure may be largely preserved in web-based adaptations of patient-reported outcomes including minor to moderate modifications (as defined by ISPOR recommendations [25]) when tested in a BD sample.

Although the EFA results support that, overall, the QoL Tool and QoL BD have concordant factor structures, one point of divergence warrants further discussion. In the QoL Tool sample, EFA results suggest one latent factor may best account for items from both the sleep and physical QoL domains, whereas the original validation study of the QoL.BD instrument supported two distinct factors. This design does not allow us to unpack the determinant of this difference. We can speculate that, potentially, the minor-to-moderate modifications in the web-based delivery of the QoL Tool (changes to wording of one item, a graphical representation of results, and the ability for users to see the domain names) may have contributed to this small divergence in factor structure. A second candidate explanation that must be considered is differing sample compositions [45]: In this study, similar proportions of individuals reporting BD-I and BD-II diagnoses participated; the original validation study predominantly consisted of individuals with BD-I (see Table 1). Given that the prevalence of various physical health comorbidities [46,47] and the experience of sleep disturbances [48] may vary according to BD subtype, heterogeneity between samples may underpin the differences in factor structure observed over the sleep and physical items viz the original validation study.

There are no clear-cut guidelines regarding the optimal way to respond when faced with points of divergence in the psychometric properties of web-based and pen-and-paper instruments; developers must consider the impact on psychometric properties in light of the numerous advantages of web-based adaptations (discussed above) as well as supporting interpretation of results by preserving the surface structure and face validity of the questionnaire. ISPOR recommendations highlight that fidelity to the original pen-and-paper instrument must be balanced against the potential to improve functionality and performance in web-based adaptations [25], and as such, concrete, universal recommendations about standards of evidence and quantitative cutoffs for acceptable psychometric properties cannot be made. In fact, the meaning of divergence between pen-and-paper instruments and web-based adaptations is not clear cut and may in fact reflect *improved* data quality and user-friendliness on the part of the web-based instrument, as the potential for social desirability effects or missing data to bias findings is ameliorated. Furthermore, psychometric findings about factor structure is only one piece of evidence which should drive decisions about the surface structure of an instrument. In the case of this study, although it is perhaps unsurprising that a latent factor may underpin items assessing both sleep and physical health, given that some of the physical health items (eg, "I have had plenty of energy") are likely to be impacted by achieving adequate sleep, there is also evidence to suggest the face validity of distinct domains. Assessing these domains separately is key for the instrument to have clinical and research utility, given that sleep changes in BD are one of the most prominent prognostic indicators of mood destabilization [49], and sleep difficulties require different self-management and clinical interventions relative to physical health comorbidities [50]. In light of this and given strong conceptual arguments for separate sleep and physical domains, we suggest these items continue to be scored according to the original QoL.BD.

Limitations

A number of limitations to this study should be noted. First, participants self-reported their diagnosis of BD; although the confirmation of diagnosis by structured psychiatric interview would have been preferable, there is some evidence that people who self-identify as living with BD typically do meet diagnostic criteria [51]. Furthermore, as the sample was self-selected, higher levels of digital literacy may have been present: qualitative interviews with a small subsample of participants who were given the opportunity to use a web-based BD self-management intervention suggest that some participants struggled to access that website because of technological barriers [34]. Care must be taken to evaluate the feasibility and psychometric properties of the QoL Tool in samples with lower levels of digital literacy or those facing a digital divide.

Finally, the web-based and pen-and-paper versions of the QoL.BD instrument were not directly compared in the same sample. Therefore, we are unable to determine whether factor equivalence was impacted by differing demographic compositions, rather than modifications to the delivery of the instrument itself. Furthermore, we were not able to assess concordance with the pen-and-paper QoL.BD in the form of intraclass correlation coefficients. However, it has been noted that it is not typically feasible nor warranted to assess test-retest reliability across instrument modes [25] and indeed, this may introduce confounding learning effects.

Implications and Future Directions

This study provides evidence for concordance between the web-based and paper-based versions of the widely adopted QoL.BD questionnaire, providing some confidence in the use of the QoL Tool in research or clinical assessments. Furthermore, there is now also qualitative evidence to suggest that the QoL Tool can be integrated positively into the self-management practices of individuals with BD; respondents described the breadth of areas assessed as enabling them to identify areas of strengths as well as areas in need of improvement [52]. This emerging body of evidence for the

utility of the QoL Tool, both from a psychometric and subjective perspective, suggests further development and dissemination of this measure is warranted.

One avenue of expansion is the translation of the QoL Tool from a web-based interface to a mobile phone app. The project to develop the QoL Tool was initiated in 2014; significant advances have occurred in the digital mental health landscape since that time—we now need to avail of developing technologies to enhance the delivery and functionality of this instrument. People with BD have shown interest in digitally supported self-management [32,53], and self-monitoring apps have been found to be feasible and acceptable in this population [54-56]. However, current apps do not adequately meet consumer needs to track a broad spectrum of outcomes [32], with most apps developed to assess domains of well-being in isolation, such as sleep or mood. Individuals with BD have described resorting to elaborate, self-generated systems to track multiple indicators [57]; between-app integration is a requested feature of apps for BD [32]. The range of wellness outcomes assessed by the QoL Tool means that its adaptation into an app format is likely to meet this consumer need.

Moving forward, CREST.BD has now initiated a 3-year project to incorporate the evidence and tools held in the Bipolar Wellness Centre [58] and the QoL Tool [19] into a new mobile health app—"Bipolar Bridges." The project aims to address some of the limitations of existing BD apps by using CBPR approaches to co-design an app that synergistically combines different forms of digital health data (including QoL Tool results), enabling individuals to learn what self-management strategies are most effective for optimizing their QoL.

Conclusions

This study provides initial support for the psychometric validity of the QoL Tool, a web-based adaptation of an instrument to measure QoL in BD (the QoL.BD instrument). Specifically, internal consistency estimates, distribution of scores, and factor structure were largely consistent with the pen-and-paper version. Although evidence supporting the overall equivalence of these instruments was observed, the findings of this study suggested a latent structure of 11 compared with the original 12 basic domains in the original instrument. This 11-factor structure combined items from the sleep and physical domains in a single shared factor. Two explanations for this minor divergence must be considered: changes to the user experience in the web-based interface and differences in sample composition. As the design of this study does not permit separating the influence of these two potential explanations, further research is required. However, in light of the face validity and clinical utility of a distinct sleep domain in QoL in BD, we recommend these items continue to be treated according to the structure of the original QoL.BD. Given the increasing role of web-based self-report questionnaires for research, clinical, and self-management contexts, findings of overall psychometric equivalence between these QoL instruments validate current applications of the QoL Tool and encourage further efforts to optimize its web-based delivery and associated self-management strategies via a novel mobile phone app.

Acknowledgments

This study was supported by a Canadian Institutes of Health Research Knowledge to Action grant, grant #201210KAL-259464. The authors would like to thank the research participants who were involved in this project, the expertise of the CREST.BD Community Advisory Group, and CREST.BD network members.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Primary factor loadings of the quality of life tool based on an exploratory factor analysis with maximum likelihood extraction and oblique rotation.
[DOCX File , 16 KB - mental_v7i4e17497_app1.docx]

References

1. Morton E, Michalak EE, Murray G. What does quality of life refer to in bipolar disorders research? A systematic review of the construct's definition, usage and measurement. J Affect Disord 2017 Apr 1;212:128-137. [doi: 10.1016/j.jad.2017.01.026] [Medline: 28160685]
2. Murray G, Michalak EE. The quality of life construct in bipolar disorder research and practice: past, present, and possible futures. Bipolar Disord 2012 Dec;14(8):793-796. [doi: 10.1111/bdi.12016] [Medline: 23131090]
3. Basu D. Quality-of-life issues in mental health care: past, present, and future. German J Psychiatry 2004;7(3):35-43 [FREE Full text]
4. Maczka G, Siwek M, Skalski M, Dudek D. [Patients' and doctors' attitudes towards bipolar disorder--do we share our beliefs?]. Psychiatr Pol 2009;43(3):301-312. [Medline: 19725423]
5. Joyce CR, O'Boyle CA, McGee H. Introduction: The individual perspective. In: Joyce CR, O'Boyle CA, McGee H, editors. Individual Quality of Life: Approaches to Conceptualisation and Assessment. New York: Routledge; 1999:3-8.

6. Bowling A. What things are important in people's lives? A survey of the public's judgements to inform scales of health related quality of life. Soc Sci Med 1995 Nov;41(10):1447-1462. [doi: 10.1016/0277-9536(95)00113-l] [Medline: 8560313]
7. Michalak EE, Murray G, Collaborative RESearch Team to Study Psychosocial Issues in Bipolar Disorder (CREST.BD). Development of the QoL.BD: a disorder-specific scale to assess quality of life in bipolar disorder. Bipolar Disord 2010 Nov;12(7):727-740. [doi: 10.1111/j.1399-5618.2010.00865.x] [Medline: 21040290]
8. Modabbernia A, Yaghoubidoust M, Lin C, Fridlund B, Michalak EE, Murray G, et al. Quality of life in Iranian patients with bipolar disorder: a psychometric study of the Persian Brief Quality of Life in Bipolar Disorder (QoL.BD). Qual Life Res 2016 Jul;25(7):1835-1844. [doi: 10.1007/s11136-015-1223-0] [Medline: 26714698]
9. Xiao L, Gao Y, Zhang L, Chen P, Sun X, Tang S. Validity and reliability of the Brief version of Quality of Life in Bipolar Disorder' (Bref QoL.BD) among Chinese bipolar patients. J Affect Disord 2016 Mar 15;193:66-72. [doi: 10.1016/j.jad.2015.12.074] [Medline: 26766036]
10. Morgado C, Tapia T, Ivanovic-Zuvic F, Antivilo A. Evaluación de la calidad de vida de pacientes bipolares chilenos: propiedades psicométricas y utilidad diagnóstica de la versión chilena del cuestionario Quality of Life Bipolar Disorder (QoL.BD-CL). Rev méd Chile 2015 Feb;143(2):213-222. [doi: 10.4067/s0034-98872015000200009]
11. CREST.BD. 2016. Quality of Life: Researching and Exchanging Knowledge on Quality of Life in Bipolar Disorder URL: http://www.crestbd.ca/research/research-areas/quality-of-life/ [accessed 2019-11-27]
12. Jones SH, Smith G, Mulligan LD, Lobban F, Law H, Dunn G, et al. Recovery-focused cognitive-behavioural therapy for recent-onset bipolar disorder: randomised controlled pilot trial. Br J Psychiatry 2015 Jan;206(1):58-66. [doi: 10.1192/bjp.bp.113.141259] [Medline: 25213157]
13. Jones SH, Knowles D, Tyler E, Holland F, Peters S, Lobban F, et al. The feasibility and acceptability of a novel anxiety in bipolar disorder intervention compared to treatment as usual: A randomized controlled trial. Depress Anxiety 2018 Oct;35(10):953-965. [doi: 10.1002/da.22781] [Medline: 30024639]
14. Beck AK, Baker A, Jones S, Lobban F, Kay-Lambkin F, Attia J, et al. Exploring the feasibility and acceptability of a recovery-focused group therapy intervention for adults with bipolar disorder: trial protocol. BMJ Open 2018 Jan 31;8(1):e019203 [FREE Full text] [doi: 10.1136/bmjopen-2017-019203] [Medline: 29391366]
15. Tyler E, Lobban F, Sutton C, Depp C, Johnson S, Laidlaw K, et al. Feasibility randomised controlled trial of Recovery-focused Cognitive Behavioural Therapy for Older Adults with bipolar disorder (RfCBT-OA): study protocol. BMJ Open 2016 Mar 3;6(3):e010590 [FREE Full text] [doi: 10.1136/bmjopen-2015-010590] [Medline: 26940112]
16. Fletcher K, Foley F, Thomas N, Michalak E, Berk L, Berk M, et al. Web-based intervention to improve quality of life in late stage bipolar disorder (ORBIT): randomised controlled trial protocol. BMC Psychiatry 2018 Jul 13;18(1):221 [FREE Full text] [doi: 10.1186/s12888-018-1805-9] [Medline: 30001704]
17. Calabrese JR, Sanchez R, Jin N, Amatniek J, Cox K, Johnson B, et al. Symptoms and functioning with aripiprazole once-monthly injection as maintenance treatment for bipolar I disorder. J Affect Disord 2018 Feb;227:649-656. [doi: 10.1016/j.jad.2017.10.035] [Medline: 29174738]
18. Yatham LN, Mackala S, Basivireddy J, Ahn S, Walji N, Hu C, et al. Lurasidone versus treatment as usual for cognitive impairment in euthymic patients with bipolar I disorder: a randomised, open-label, pilot study. Lancet Psychiatry 2017 Mar;4(3):208-217. [doi: 10.1016/S2215-0366(17)30046-9] [Medline: 28185899]
19. Michalak E, Murray G. CREST.BD Quality of Life Tool. 2015. URL: https://www.bdqol.com/ [accessed 2019-11-27]
20. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. J Med Internet Res 2007 Sep 30;9(3):e25 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]
21. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298. [doi: 10.1093/aje/kwq291] [Medline: 20880962]
22. Miller ET, Neal DJ, Roberts LJ, Baer JS, Cressler SO, Metrik J, et al. Test-retest reliability of alcohol measures: is there a difference between internet-based assessment and traditional methods? Psychol Addict Behav 2002 Mar;16(1):56-63. [Medline: 11934087]
23. Wijndaele K, Matton L, Duvigneaud N, Lefevre J, Duquet W, Thomis M, et al. Reliability, equivalence and respondent preference of computerized versus paper-and-pencil mental health questionnaires. Comput Hum Behav 2007;23(4):1958-1970 [FREE Full text] [doi: 10.1016/j.chb.2006.02.005]
24. Richman WL, Kiesler S, Weisband S, Drasgow F. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. J Appl Psychol 1999;84(5):754-775. [doi: 10.1037/0021-9010.84.5.754]
25. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, ISPOR ePRO Task Force. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429 [FREE Full text] [doi: 10.1111/j.1524-4733.2008.00470.x] [Medline: 19900250]
26. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity

for the Use of Existing Instruments and Their Modification PRO Task Force Report. Value Health 2009;12(8):1075-1083 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00603.x] [Medline: 19804437]
27. Alfonsson S, Maathz P, Hursti T. Interformat reliability of digital psychiatric self-report questionnaires: a systematic review. J Med Internet Res 2014 Dec 3;16(12):e268 [FREE Full text] [doi: 10.2196/jmir.3395] [Medline: 25472463]
28. van Ballegooijen W, Riper H, Cuijpers P, van Oppen P, Smit JH. Validation of online psychometric instruments for common mental health disorders: a systematic review. BMC Psychiatry 2016 Feb 25;16:45 [FREE Full text] [doi: 10.1186/s12888-016-0735-7] [Medline: 26915661]
29. Satcher D, Israel BA, Eng EE, Schulz AJ, Parker EA. Methods in Community-Based Participatory Research for Health. San-Fransico, CA: Jossey-Bass; 2005.
30. Michalak E, Lane K, Hole R, Barnes S, Khatri N, Lapsley S, et al. Towards a better future for Canadians with bipolar disorder: principles and implementation of a community-based participatory research model. Engaged Scholar J 2015;1(1):132-147 [FREE Full text] [doi: 10.15402/esj.2015.1.a08]
31. Woods CM. Careless responding to reverse-worded items: implications for confirmatory factor analysis. J Psychopathol Behav Assess 2006;28(3):186-191. [doi: 10.1007/s10862-005-9004-7]
32. Nicholas J, Fogarty AS, Boydell K, Christensen H. The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. J Med Internet Res 2017 Apr 7;19(4):e105 [FREE Full text] [doi: 10.2196/jmir.7273] [Medline: 28389420]
33. 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]
34. Michalak EE, Morton E, Barnes SJ, Hole R, CREST.BD, Murray G. Supporting self-management in bipolar disorder: mixed-methods knowledge translation study. JMIR Ment Health 2019 Apr 15;6(4):e13493 [FREE Full text] [doi: 10.2196/13493] [Medline: 30985287]
35. Tabachnick BG, Fidell LS. Using Multivariate Statistics. Boston, MA: Pearson; 2007.
36. Nunnally JC, Bernstein IH. Psychometric Theory. New York: McGraw-Hill; 1994.
37. Williams B, Onsman A, Brown T. Exploratory factor analysis: A five-step guide for novices. Australas J Paramedicine 2010;8(3):1-13. [doi: 10.33151/ajp.8.3.93]
38. Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 1960;20(1):141-151. [doi: 10.1177/001316446002000116]
39. O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. Behav Res Methods Instrum Comput 2000 Aug;32(3):396-402. [doi: 10.3758/bf03200807] [Medline: 11029811]
40. Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika 1976;41(3):321-327. [doi: 10.1007/bf02293557]
41. Pett MA, Lackey NR, Sullivan JJ. Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research. Thousand Oaks, CA: Sage; 2003.
42. Howard MC. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? Int J Hum-Comput Interact 2016;32(1):51-62. [doi: 10.1080/10447318.2015.1087664]
43. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod 2013 Feb;38(1):52-54 [FREE Full text] [doi: 10.5395/rde.2013.38.1.52] [Medline: 23495371]
44. Comrey AL, Lee HB. A First Course in Factor Analysis. Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1992.
45. Herrero J, Meneses J. Short Web-based versions of the perceived stress (PSS) and Center for Epidemiological Studies-Depression (CESD) Scales: a comparison to pencil and paper responses among Internet users. Comput Hum Behav 2006;22(5):830-846 [FREE Full text] [doi: 10.1016/j.chb.2004.03.007]
46. Kemp DE, Sylvia LG, Calabrese JR, Nierenberg AA, Thase ME, Reilly-Harrington NA, LiTMUS Study Group. General medical burden in bipolar disorder: findings from the LiTMUS comparative effectiveness trial. Acta Psychiatr Scand 2014 Jan;129(1):24-34 [FREE Full text] [doi: 10.1111/acps.12101] [Medline: 23465084]
47. Forty L, Ulanova A, Jones L, Jones I, Gordon-Smith K, Fraser C, et al. Comorbid medical illness in bipolar disorder. Br J Psychiatry 2014 Dec;205(6):465-472 [FREE Full text] [doi: 10.1192/bjp.bp.114.152249] [Medline: 25359927]
48. Lewis KS, Gordon-Smith K, Forty L, Di Florio A, Craddock N, Jones L, et al. Sleep loss as a trigger of mood episodes in bipolar disorder: individual differences based on diagnostic subtype and gender. Br J Psychiatry 2017 Sep;211(3):169-174 [FREE Full text] [doi: 10.1192/bjp.bp.117.202259] [Medline: 28684405]
49. Morton E, Murray G. An update on sleep in bipolar disorders: presentation, comorbidities, temporal relationships and treatment. Curr Opin Psychol 2019 Aug 26;34:1-6. [doi: 10.1016/j.copsyc.2019.08.022] [Medline: 31521023]
50. Morton E, Murray G. Assessment and treatment of sleep problems in bipolar disorder-A guide for psychologists and clinically focused review. Clin Psychol Psychother 2020 Feb 18 [Epub ahead of print]. [doi: 10.1002/cpp.2433] [Medline: 32068325]

51. Kupfer DJ, Frank E, Grochocinski VJ, Cluss PA, Houck PR, Stapf DA. Demographic and clinical characteristics of individuals in a bipolar disorder case registry. J Clin Psychiatry 2002 Feb;63(2):120-125. [doi: 10.4088/jcp.v63n0206] [Medline: 11874212]
52. Morton E, Hole R, Murray G, Buzwell S, Michalak E. Experiences of a web-based quality of life self-monitoring tool for individuals with bipolar disorder: a qualitative exploration. JMIR Ment Health 2019 Dec 4;6(12):e16121 [FREE Full text] [doi: 10.2196/16121] [Medline: 31799936]
53. 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]
54. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]
55. Hidalgo-Mazzei D, Mateu A, Reinares M, Murru A, del Mar Bonnín C, Varo C, et al. Psychoeducation in bipolar disorder with a SIMPLe smartphone application: feasibility, acceptability and satisfaction. J Affect Disord 2016 Aug;200:58-66. [doi: 10.1016/j.jad.2016.04.042] [Medline: 27128358]
56. Schwartz S, Schultz S, Reider A, Saunders EF. Daily mood monitoring of symptoms using smartphones in bipolar disorder: a pilot study assessing the feasibility of ecological momentary assessment. J Affect Disord 2016 Feb;191:88-93 [FREE Full text] [doi: 10.1016/j.jad.2015.11.013] [Medline: 26655117]
57. 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]
58. CREST.BD. 2015. Bipolar Wellness Centre URL: http://www.crestbd.ca/tools/bipolar-wellness-centre/ [accessed 2019-11-27]

Abbreviations

BD: bipolar disorder
BD-I: bipolar disorder type 1
BD-II: bipolar disorder type 2
CBPR: community-based participatory research
CREST.BD: The Collaborative RESearch Team to study psychosocial issues in Bipolar Disorder
EFA: exploratory factor analysis
ISPOR: Professional Society for Health Economics and Outcomes Research
KT: knowledge translation
MAP: minimum average partial correlation
NOS: not otherwise specified
QoL.BD: Quality of Life in Bipolar Disorder
QoL: quality of life

Edited by J Torous; submitted 17.12.19; peer-reviewed by B Johnson; comments to author 03.02.20; accepted 09.02.20; published 27.04.20.

Please cite as:
Morton E, Hou SHJ, Fogarty O, Murray G, Barnes S, Depp C, CREST.BD, Michalak E
A Web-Based Adaptation of the Quality of Life in Bipolar Disorder Questionnaire: Psychometric Evaluation Study
JMIR Ment Health 2020;7(4):e17497
URL: http://mental.jmir.org/2020/4/e17497/
doi:10.2196/17497
PMID:32338620

Review

The Use of Text Messaging to Improve Clinical Engagement for Individuals With Psychosis: Systematic Review

Jessica D'Arcey[1,2], BSc; Joanna Collaton[3], MPH; Nicole Kozloff[1,4], MD, SM; Aristotle N Voineskos[1,2,5], MD, PhD; Sean A Kidd[1,2], PhD; George Foussias[1,2,5], MD, PhD

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

[2]Institute of Medial Science, University of Toronto, Toronto, ON, Canada

[3]Department of Clinical Psychology, University of Guelph, Guelph, ON, Canada

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

[5]Department of Psychiatry, Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
George Foussias, MD, PhD
Slaight Centre for Youth in Transition
Centre for Addiction and Mental Health
250 College St 738
Toronto, ON, M5T1R8
Canada
Phone: 1 4165358501 ext 34390
Email: george.foussias@camh.ca

Abstract

Background: Individuals experiencing psychosis are at a disproportionate risk for premature disengagement from clinical treatment. Barriers to clinical engagement typically result from funding constraints causing limited access to and flexibility in services. Digital strategies, such as SMS text messaging, offer a low-cost alternative to potentially improve engagement. However, little is known about the efficacy of SMS text messaging in psychosis.

Objective: This review aimed to address this gap, providing insights into the relationship between SMS text messaging and clinical engagement in the treatment of psychosis.

Methods: Studies examining SMS text messaging as an engagement strategy in the treatment of psychosis were reviewed. Included studies were published from the year 2000 onward in the English language, with no methodological restrictions, and were identified using 3 core databases and gray literature sources.

Results: Of the 233 studies extracted, 15 were eligible for inclusion. Most studies demonstrated the positive effects of SMS text messaging on dimensions of engagement such as medication adherence, clinic attendance, and therapeutic alliance. Studies examining the feasibility of SMS text messaging interventions found that they are safe, easy to use, and positively received.

Conclusions: Overall, SMS text messaging is a low-cost, practical method of improving engagement in the treatment of psychosis, although efficacy may vary by symptomology and personal characteristics. Cost-effectiveness and safety considerations were not adequately examined in the studies included. Future studies should consider personalizing SMS text messaging interventions and include cost and safety analyses to appraise readiness for implementation.

(JMIR Ment Health 2020;7(4):e16993) doi:10.2196/16993

KEYWORDS

SMS; text messaging; psychosis; schizophrenia; bipolar disorder; engagement; medication adherence; attendance; patient appointments

Introduction

Psychosis and Engagement

The major psychoses (ie, schizophrenia spectrum and bipolar I disorder) are recognized as a leading cause of disability worldwide [1] and are associated with poverty [2], premature mortality [3], impaired cognitive function [4], loss of education and employment [5], and increased global economic burden [6]. Published practice guidelines outlining the usual treatment for psychosis identify antipsychotic medication as a frontline treatment for positive symptom management and adjunct psychosocial interventions such as psychoeducation, family support, vocational interventions, and cognitive behavioral skills–based therapies [7-11]. However, despite concerted treatment efforts, disengagement from clinical services remains to be a significant barrier to recovery for this population as evidenced by high rates of nonadherence (25%-50%) [12], nonattendance (40%) [13], and service dropout (30%) [14].

Background

Clinical engagement is a combination of health-oriented attitudes (ie, beliefs about the cause of illness and intentions of the care team) and behaviors (ie, appointment attendance, adherence to treatment plans, and the level of participation in a clinical relationship) [15,16]. Therefore, disengagement occurs when attitudinal factors cause patients to stop performing treatment behaviors such as taking medications and attending appointments [16]. Disengaged individuals often exhibit greater functional impairment and higher symptom burden over time [17,18], subsequently leading to higher rates of relapse and rehospitalization [12,13]. Cycles of relapse and rehospitalization create a unique conundrum for this population as evidence suggests that with each period of disengagement and relapse, positive symptoms become more severe and more difficult to treat [19]. This group is also less likely to fully recover or hold employment [18]. Moreover, studies show that early and sustained engagement in treatment can mitigate negative outcomes of psychosis such as significantly lowering all-cause mortality [20], reducing the risk of relapse [7-11] and ameliorating functional impairment [5,9]. As such, understanding and improving engagement in this population is a priority.

Barriers to clinical engagement are rooted in access to treatment including geographic, systemic, and financial obstacles. Psychosocial services are often located in urbanized areas, serving specific catchment areas [21] and limiting access for rural or remote residents [15,21-23]. Moreover, services are often only available in a business model (ie, Monday-Friday, 9 AM-5 PM) [24-26], with clinicians often treating double the recommended caseloads, further restricting appointment availability [26] because of funding constraints [24-26]. These challenges are often exacerbated by personal factors such as poverty, limited transportation, substance use, trauma history, and beliefs about mental health [15].

Given the financial and staff limitations, there has been a call for more cost-effective, flexible modes of engagement, such as SMS text messaging. SMS text messaging is a globally used mechanism for communication across both wealthy and poor countries [27], and cellular phone ownership and its use in populations with a psychiatric illness is high (72%-97%) [28-31]. SMS text messaging offers an easy, user-friendly platform to extend opportunities for engagement beyond clinical boundaries. SMS text messaging also does not require access to a smartphone, wireless internet, or cellular data, making it a more accessible platform for low-income populations, including those with psychosis. These characteristics may allow SMS text messaging to provide support remotely, and in real time, to boost engagement as shown in other areas of health care such as HIV [32], diabetes [33-35], coronary heart disease [36], obesity [37], and substance use disorders [38,39].

This Study

This systematic review aimed to examine studies investigating SMS text messaging strategies to improve clinical engagement in individuals with psychosis. To date, reviews have focused on the subdomains of clinical engagement such as clinic attendance or medication adherence; however, no such review or single study has examined engagement holistically nor have reviews focused on the use of SMS text messaging within populations with psychosis. This is an important knowledge gap given the challenging nature of engagement in the treatment of psychosis, the resulting considerable outcome disparities, and recent trends toward technologically aided health care. As such, even a modest improvement in clinical engagement could lead to better cost-effectiveness of psychosis treatments [40].

Methods

Study Inclusion Criteria

Studies were included regardless of the methodological framework or quality, patient setting (ie, inpatient, outpatient, or community care), age of the population, or stage of illness to provide an inclusive examination of clinical engagement. Studies with a range of quality ratings were included to fully examine the state of the literature on this topic. Included interventions used SMS text messaging as a delivery platform, whereas interventions with mixed or non–SMS text messaging technology (eg, mobile apps) were excluded. This review focused on treatment-seeking individuals with psychosis (ie, the portion of the population either beginning or receiving treatment) to isolate the effect of SMS text messaging on clinical engagement. Although it is true that disengagement is a challenge for those who have yet to be in contact with the mental health care system, attempting to study this population would shift focus from treatment engagement to treatment outreach.

Primary Outcomes: Clinical Engagement

Clinical engagement is a broad term describing patients' participation in treatment and is typically measured using its behavioral outcomes such as appointment attendance (eg, number or percentage of appointments attended, cancelled, or rescheduled), medication adherence (eg, pill counts, clinician or self-estimates, pharmacy prescriptions, or blood plasma levels), and service dropout (eg, complete disengagement from care). However, the consideration of behavioral outcomes alone does not provide a holistic examination of clinical engagement as it excludes the attitudinal aspects of this complex and nuanced

concept. To this end, we also included studies with outcomes measuring attitudes toward medication, the level of involvement or participation in treatment plans, and therapeutic rapport.

Secondary Outcomes: Feasibility

Feasibility studies were also included to assess 2 outcomes: the level of patient engagement with the SMS text messaging intervention itself, and the degree of practicality of SMS text messaging interventions and their ease of implementation. Outcomes of feasibility trials include user feedback on ease of use, satisfaction, likelihood of future use, and overall intervention design. Outcomes related to practicality and implementation include associated risk and economic, technical, and legal considerations. These are important outcomes to consider when assessing the efficacy of interventions because both directly affect the interventions' clinical utility. An intervention of any kind is only as successful as its ability to engage its users. In this case, engagement in the SMS text messaging interventions aims to improve engagement in the treatment of psychosis. Thus, this review examined both aspects of engagement to comprehensively evaluate the efficacy and clinical utility of SMS text messaging interventions to improve clinical engagement in populations with psychosis.

Search Strategy

Web-based databases cataloging meta-analyses, systematic reviews, and protocols were searched to ensure the originality of this protocol (ie, Cochrane and PROSPERO); however, there were no results on existing reviews on the use of SMS text messaging in psychosis. A unique protocol was, then, created following the Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines and in consult with a librarian at the University of Toronto. The resulting protocol was registered with PROSPERO (reference number: CRD42018091962), an international protocol registry for prospective reviews in health care managed by the National Institute of Health Research.

In all, 3 core databases powered by OVID were used: PsycINFO, MEDLINE (Medical Literature Analysis and Retrieval System Online), and EMBASE, using the following search strategy: schizophrenia spectrum.mp. OR psychotic disorder.mp. OR exp psychosis/ OR exp bipolar disorder/AND (sms or short messag* service* or texting or text messag*).mp. Please note that *.mp.* is a mapping command that allows you to search using a word or phrase across titles, keywords, and abstracts using databases powered by OVID.

CINAHL and Google Scholar were used as peripheral databases to ensure studies were not missed. Search terms and keywords related to psychosis and SMS text messaging that were used to search gray literature included the following: (1) SMS OR Short Message Service OR SMS-Survey OR Texting OR Text Message OR SMS Based System OR SMS Reminder OR Text Message Reminder and (2) Psychosis OR Schizophrenia OR Bipolar OR Schizoaffective OR Schizophrenia Spectrum OR Psychotic Disorders OR First-Episode Psychosis OR Early-Episode Psychosis. Additionally, reference lists of identified studies were hand searched.

The search criteria were limited to studies published in the English language, restricting publication dates from January 2000 to March 2019 as there were no SMS text messaging intervention studies published before the year 2000.

Review Process

Studies were independently reviewed by title, abstract, or full-text using a Web-based blinded review platform named Covidence, which allows the reviewers to log-in independently to conduct their review and tracks conflicts to be resolved. Conflicts were resolved first by the 2 primary reviewers, authors JD and JC, and then by a third reviewer, GF, if necessary. Only full-length empirical studies were reviewed; conference abstracts and commentaries were excluded given their limited descriptions of methodologies and inability to be adequately appraised.

Data Extracted

We extracted sample sizes, and demographic characteristics including age and gender, and diagnoses. Methods were also extracted, including the study design, intervention type, primary and secondary outcome variables, and statistical plan. To better understand the significance of the results, effect sizes were also extracted from studies with randomized controlled trial (RCT) and quasi-experimental designs and converted into a comparable measure of effect size, Cohen *d*. Finally, results reported in each study were extracted, including primary and secondary results (including *P* values, if available) as well as any other findings of interest to the topic of the review.

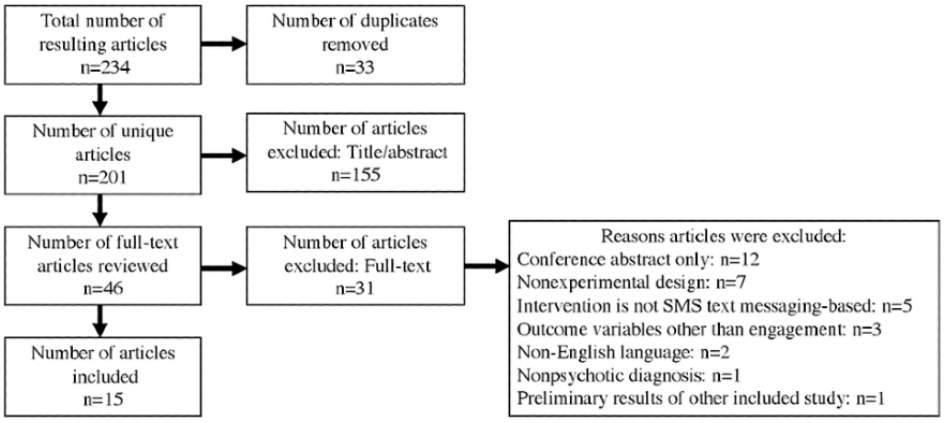
Quality Appraisal

As studies were included regardless of the study design, 3 checklists from the Joanna Briggs toolkit were used: RCTs, quasi-experimental designs, and qualitative studies [41]. This toolkit is an open-access, widely used appraisal tool for systematic reviews accessible through the citation provided. Using the checklist that corresponded to the study type, studies received 1 point per criterion met. Points were tallied and converted to a percentage based on the total number of criteria outlaid on the checklist used for easy comparison across study designs. However, results were compared and contrasted based on the matching outcome variable and study design.

Results

The original searches of databases yielded 234 studies (Figure 1). The removal of duplicates, as well as a review of titles and abstracts, led to the exclusion of 188 studies. A subsequent full-text review of the remaining 46 articles led to the exclusion of an additional 31 articles, resulting in 15 studies meeting the criteria for inclusion in this review: 8 RCTs, 1 quasi-experimental design, and 6 qualitative studies.

Figure 1. Flow chart of search results.



Study Characteristics

As outlined in Table 1, studies were conducted in a range of countries, although most were Western, developed countries from Europe and North America, with 1 study each from Nigeria, China, and India. They were published from 2010 to 2018, with steady publication rates throughout. Most studies included patients with a diagnosis of schizophrenia spectrum disorders, aged 18 years and older, with 1 study focusing on a younger population, aged 18 to 29 years. Sample sizes ranged from 17 to 1139 with some studies based on the same sample but reporting on different outcomes (eg, studies by Ben-Zeev [42] and Aschbrenner et al [43]; and studies by Kauppi et al [44], Kannisto et al [45], and Välimäki et al [46]). RCTs tended to examine the effects of SMS text messaging reminders on medication adherence and/or clinic attendance paired with secondary outcomes related to symptom severity and functioning. Qualitative studies examined the feasibility and acceptability of SMS text messaging interventions.

Table 1. Characteristics of the studies included.

References	Year	Country	Diagnosis/population	Values, n	Age (years), range (mean)	Male:Female
Välimäki et al [46]	2017	Finland	Psychosis	1139[a]	18-65 (38.3)	1:2
Kauppi et al [44]	2015	Finland	Psychosis	562[a]	18-65 (38.6)	1:1
Kannisto et al [45]	2015	Finland	Psychosis	403[a]	18-65 (39.7)	1:1
Montes et al [47]	2012	Spain	SZ[b]	340	18-65 (39.6)	2:1
Xu [48]	2017	China	SZ	237	35-60 (45)	1:1
Menon et al [49]	2018	India	Bipolar I disorder	132	18-65 (37.9)	1:1
Beebe et al [50]	2014	United States	SZ and SZA[c]	30	21-68 (48.7)	1:2
Thomas et al [51]	2017	Nigeria	Psychosis	192	18-64 (33.7)	1:1
Pijnenborg et al [52]	2010	Netherlands	SZ spectrum	62	NR[d] (28)	4:1
Kravarti et al [53]	2018	United Kingdom	SZ spectrum	75	NR (42.14)	1:1
Granholm et al [54]	2012	United States	SZ and SZA	55	18 (48)	2:1
Ben-Zeev et al [42]	2014	United States	SZ, SZA, and SU[e]	28[f]	18 (40.5)	2:1
Aschbrenner et al [43]	2016	United States	SZ, SZA, and SU	17[f]	18 (NR)	NR
Lal et al [55]	2015	Canada	Early psychosis	67	18-35 (25.6)	3:1
Bogart et al [56]	2014	United Kingdom	Antipsychotic use	85	18-67 (NR)	1:1

[a]Indicates a shared sample.

[b]SZ: schizophrenia.

[c]SZA: schizoaffective disorder.

[d]NR: not reported.

[e]SU: substance use.

[f]Indicates a shared sample.

Appraisal of Study Quality

Studies included in this review were of a reasonable quality (range 31%-85%); however, no study met all of the criteria laid out in the Joanna Briggs frameworks. Individual study ratings are included in Table 2. RCTs received a range of scores from 31% to 85%, with sample sizes between 28 and 1119 participants. Common unmet criteria in RCTs were blinding expectations, as none of the included trials were double-blinded and many failed to use a concealed assignment. Additionally, some trials were underpowered for their results. In the quasi-experimental study, there was no control group used, causing multiple criteria to be unmet. Qualitative designs tended to satisfy a higher proportion of the quality criteria, receiving scores from 40% to 80% with sample sizes ranging from 17 to 562. These studies often lacked information around experimenter bias and sampling methods and an adequate representation of the population. Of note, feasibility studies do not fit smoothly within the qualitative study category and, therefore, appraisal was limited in this domain. Despite these methodological limitations, these studies provide adequate assessments of and insights into a variety of outcomes related to SMS text messaging, clinical engagement, and the consideration of SMS text messaging implementations.

The quality of outcomes within the included studies was fair, although heterogeneous. The included studies showed inconsistent methods of measurement across medication adherence, often relying on self-report, and the calculation of attendance rates was unclear. Other considerations included small sample sizes as 4 of the experimental studies had samples under 100, and only 1 trial had insufficient power to detect a significant difference [50]. The resulting effect sizes were also variable, ranging from moderate to high. Additionally, the intervention design was inconsistent from the type of SMS text messaging (ie, one-way vs two-way messaging) to the frequency, length, and content of the messages themselves. Only 2 of the included studies were designed to include a follow-up period, both examining participants' adherence up to 3 months postintervention [47,49]. Taken together, this amount of variance makes it difficult to assess the overall success of the included interventions; however, no clear patterns emerge to suggest that sample size, message type, or effect size varies in response to any methodological factors in these trials. Individual study intervention details are outlined in Table 2 and procedures are outlined in Table 3.

A summary of the main results has been provided in Multimedia Appendix 1.

Table 2. SMS text messaging intervention details of the studies included that performed a trial.

References	Study setting	Reminder type	Frequency of SMS text messaging	Length	Follow-up?	Compensation	Delivery platform	Automated vs Manual	One- vs two-way messaging
Välimäki et al [46], Kannisto et al [45], Kauppi et al [44][a]	Inpatient discharge	SMS text messaging	2 to 12 messages per month based on the participant	1 year	No	NR[b]	Web platform	Semiautomated	One-way
Montes et al [47]	Outpatient/Community	SMS text messaging	Daily	3 months	Yes, 3 months post	NR	Web platform	Automated	One-way
Xu [48]	Rural community	SMS text messaging to the patient/lay health support person	Up to 2 messages per day based on the participant	6 months	No	Points per response in exchange for hygiene items	Web platform	Semiautomated	Two-way
Menon et al [49]	Outpatient	SMS	2 messages per week	3 months	Yes, 3 months post	NR	Cell phone	Manual (study principal investigator)	One-way
Beebe et al [50]	Community	SMS text messaging and a call	Daily	3 months	No	US $10 per monthly assessment	Cell phone	Manual (mental health nurse)	Two-way
Thomas et al [51]	Outpatient	SMS text messaging	5 and 3 days before the appointment	5 days	No	NR	Web platform	Automated	One-way
Pijnenborg [52]	Outpatient	SMS text messaging	1 message per week	7 weeks	No	NR	Web platform	NR	One-way
Kravariti [53]	Community	SMS text messaging	7 days and 1 day before the appointment	6 months	No	Not paid	Web platform	Automated	One-way
Granholm et al [54]	Outpatient	SMS text messaging	12 messages per day	3 months	No	US $35 for in-person assessments, US $20 for text message-based assessments	Web platform	Automated	Two-way
Been-Zeev et al [42], Aschbrenner et al [43][a]	Community	SMS text messaging vs calls	Up to 3 messages per day based on the preference	12 weeks	No	Reimbursed up to EUR 30 per month	Cell phone	Manual (social worker)	Two-way

[a]Indicates a shared sample.

[b]NR: not reported.

Table 3. Methodological characteristics of the study design and methods in the included studies.

References	Study design	Engagement target	Primary outcome	Measurement	Secondary outcomes	Analysis	Quality rating (%)
Studies based on intervention trials							
Välimäki [46][a]	RCT[b] 1:1 (TAU[c])	MedAd[d]	Number of Hospitalizations	Chart review	Admission type, quality of life, and user satisfaction	OR[e] and risk ratio	62
Montes et al [47]	RCT 1:1 (TAU)	MedAd	MedAd	Self-report	Symptoms, insight, quality of life, and treatment attitude	Stepwise linear regression	69
Xu [48]	RCT 1:1 (WC[f])	MedAd and AppAttd[g]	MedAd and AppAttd	Pill count and scripts	Symptoms and functioning	Generalized estimating equation	85
Menon et al [49]	RCT 1:1 (WC)	MedAd and treatment attitude	MedAd	Self-report	Symptoms and quality of life	Repeated measures ANOVA[h]	77
Beebe et al [50]	RCT 1:1:1[i] (NC)	MedAd	MedAd	Pill count	Symptoms	ANOVA	54
Thomas et al [51]	RCT 1:1 (TAU)	Initial AppAttd	Initial AppAttd	Attendance dichotomous variable (Yes or No)	Duration of untreated psychosis and symptoms	OR	69
Pijnenborg et al [52]	RCT 1:1 (WC)	AppAttd	Number of goals attained (including AppAttd)	Number of appointments attended	Role functioning, symptoms, cognition, and treatment attitude	Multiple linear regression	31
Kravariti et al [53]	RCT 1:1 (TAU)	AppAttd	AppAttd	Number of appointments Attended	N/A[j]	Proportions (%), OR	62
Granholm et al [54]	Quasi-Experimental (NCon)	MedAd	MedAd, symptoms, and socialization	Ambulatory monitoring	Role functioning and cognition	HGLM[k]	55
Studies based on feasibility trials							
Ben-Zeev et al [42][a]	Feasibility	TxAll	User engagement	Self-report	User feedback	Proportions (%) and paired *t* test	80
Aschbrenner et al [43][a]	Qualitative	Reminder interest	User interest	Thematic coding of SMS text messages	N/A	Thematic analysis	70
Lal et al [55]	Feasibility	Preferred platform	User interest	Survey	N/A	Proportions (%)	80
Bogart et al [56]	Feasibility (survey)	MedAd	User feedback	Self-report	N/A	Proportions (%) and stepwise multilinear regression	40
Kauppi et al [44][a]	Feasibility	MedAd and AppAttd	Preferred topic	Patient message selection	Preferred timing	Proportions (%)	60
Kannisto et al [45][a]	Feasibility	MedAd	User feedback	Survey	Preferred topic and platform	Proportions (%)	62

[a]Indicates a shared sample.

[b]RCT: randomized controlled trial.

[c]TAU: treatment as usual.

[d]MedAd: medication adherence.

[e]OR: odds ratio.

[f]WC: waitlist control.

[g]AppAttd: appointment adherence.

[h]ANOVA: analysis of variance.

[i]Group allocation for this study is daily SMS text messaging only, weekly phone calls only, and a combined group (daily SMS text messaging and weekly phone calls).

[j]N/A: not applicable.

[k]HGLM: Hierarchical General Linear Modelling.

Primary Outcome: Clinical Engagement

Medication Adherence

In total, 6 of the 8 experimental studies were aimed specifically at improvement of medication adherence, with most studies reporting moderate improvements [49-52]; 2 studies reported on the effect of SMS text messaging medication reminders on participatns attitude toward medication, which showed significant improvement over the course of the intervention [47,48]; 2 other studies reported on a follow-up period and found evidence of potential maintenance effects 3 months after the cessation of the SMS text messaging intervention [47,49]; and 1 study did not report positive findings using hospitalization rates as a function of medication adherence as their primary outcome [46].

In total, 3 studies reported a nonsignificant positive change in the overall group that was significant in subgroups. Specifically, efficacy was influenced by baseline adherence and participants' type of living environment such that those with low baseline adherence [48], or those living independently [54], showed significant improvement while individuals with high baseline adherence [48] or those living with a support person [48,54] showed stable and unchanged adherence rates. Additionally, 1 study even included support persons in the design of the SMS text messaging intervention to receive back up SMS text messaging reminders in the event that participants did not respond [48]. This study shows a positive change in adherence, although, unfortunately, it is not clear how much involvement was required from the individual in the support role [48].

An effectiveness trial compared weekly phone calls with a mental health nurse to daily SMS text messaging medication reminders and a combination group [50]. This study found that phone and SMS text messaging reminders showed the most improvement, followed by the phone only group, then SMS text messaging only [50]. This suggests that adherence also may be boosted with the addition of weekly clinician calls. Participants who were prescribed depot medications were included in this study and were unevenly distributed among the treatment groups (phone+SMS text messaging= 3; SMS text messaging only= 5; and phone only=7), which may bias adherence rates favorably for the phone only group [50].

One of the qualitative studies described underlying reasons for nonadherence. Patients reported that a large proportion (49%) of missed doses were purposeful, rather than owing to forgetfulness (35%) [56]. The most common reasons for intentionally skipping doses were as follows: side effects, attenuated symptoms, and thinking a lower dose would be more beneficial [56]. Negative attitudes toward medication, insufficient information about medication, and being male led to a 3 to 4 times greater likelihood of reporting purposely skipping doses of medication [56].

Appointment Attendance and Therapeutic Rapport

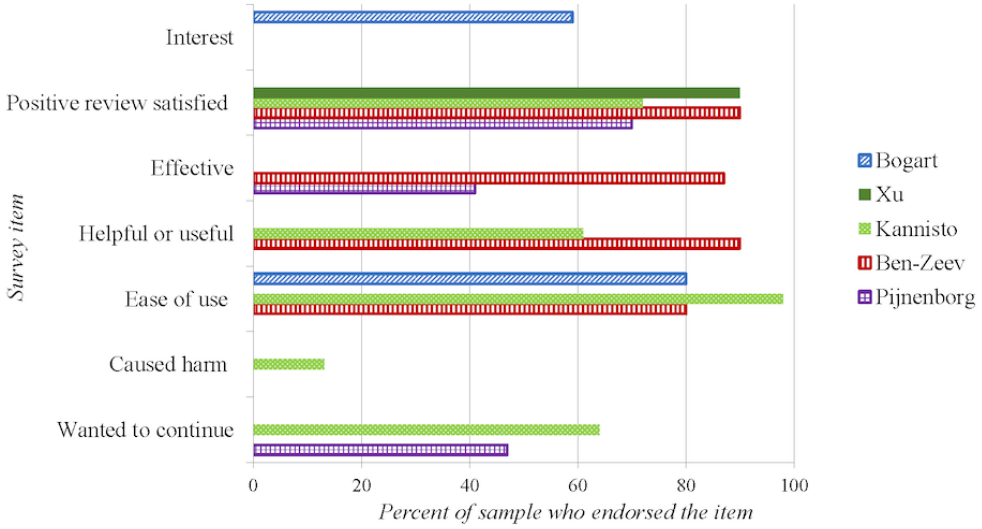
In total, 3 studies examined the efficacy of SMS text messaging reminders to improve appointment attendance. Studies demonstrated that SMS text messaging interventions are successful in increasing clinic attendance. SMS text messaging reminders increased the likelihood of attending initial appointments [51] and general attendance during the course of the intervention [52,53]. Studies comparing patient-reported rapport ratings of their usual clinical care teams, and their research-based SMS text messaging interventionist found that, typically, the SMS text messaging interventionist was rated more favorably [42,48]. This is interesting given the difference in modality of the relationship, one based in-person and the other purely electronic and text based.

Secondary Outcome: Feasibility

Practical Considerations

In total, 5 studies, including both feasibility designs and experimental designs, examined aspects of feasibility, usability, and user satisfaction as outlined in Figure 2. Studies reported a good endorsement of interest (59%) [56], wanting to continue using the intervention (47%-64%) [42,45,52], moderate-to-high ratings of effectiveness (41%-87%) [42,52], satisfaction (70%-90%) [42,45,48,52], and ease of use (80%-98%) [42,45,56]; only a small proportion endorsed harm associated with the intervention (13%) [45]. Sources of harm included participants finding the SMS text messaging reminder to be disruptive to sleep or work, or annoying [45]. Additionally, 2 studies reported on mobile phone ownership, which was high, ranging from 82% to 94% [55,56]. Other practical considerations include legal implications (eg, privacy and personal health information), work load, and cost-effectiveness. One study reported on cost, which estimated the total cost of the texting intervention from a technological and staff perspective but did not analyze how the intervention changed the overall cost in the clinic or if it was cost-efficient [46]. Another study highlighted concerns and considerations around privacy and risks associated with loss, theft, or disposal of mobile phones containing health information, recommending organizations to perform a detailed risk analysis of network security and processes for ensuring informed consent; however, no specific recommendations were made, and privacy was not surveyed or examined in the trial design [42].

Figure 2. Feasibility data collected from included studies.



Engagement With the SMS Text Messaging Intervention

The success of the interventions depended largely on the patients' level of participation and response. In all, 3 studies divided their samples using a participation threshold differentiating who tended to respond to SMS text messages and who did not [47,52,54]. The subgroup that tended not to respond to messages showed several demographic and baseline clinical differences. This group was younger, consisted of males, not working or in school, and younger at first contact with treatment [45] and were on polar ends of the clinical and functional spectrum [47,52,54]. In a minority of patients, inpatient hospitalization was reported as a barrier to answering SMS text messages as access to mobile devices is sometimes limited [42]. In the group that did tend to respond to SMS text messages, response rates were high, ranging from 85% to 87% [42,54] regardless of the message topic [54].

Generally, study attrition rates were low, ranging from 0% to 20% (see Table 2). Barriers to using SMS text messaging reported by participants included the following: 22% reported no interest in using such platforms in care, 20% reported lack of time, and 30% reported no barriers; only 7.5% reported lack of device (cell phone) as a barrier [55]. Limitations in understanding engagement with SMS text messaging are centered around the inability to confirm whether or not patients actually received the SMS text messaging reminder in the absence of a response; thus, one-way SMS text messaging cannot assess SMS text messaging participation at all outside of a participant's report; two-way messaging is still not definitive, as patients may see and not respond to the reminders, leading to reliance on the self-report as well.

User Preferences

Studies exploring participant preferences examined themes and topics of interest, as well as the frequency and timing of messages. Themes of interest to patients included messages about the following: mental health symptoms, treatment and management (eg, medication and appointments), lifestyle behaviors (eg, exercise, diet, and sleep), social relationships and leisure activity, motivation and goal setting, and independent living [43]. Another study examining personalized SMS text messaging reminders found the most requested reminders were for medication, appointments, and physical health (eg, exercise and diet) [44]. A survey study found that patients were interested in learning about a variety of psychoeducational topics including medication and side effects via digital platforms such as SMS text messaging [55].

Kauppi et al [44] examined trends in operational preferences, reporting that the preferred average number of texts per month was 10, with a preference for delivery early in the week (eg, Monday or Tuesday) and in the morning (eg, between 6 AM and noon). However, preferences differed according to a number of demographic factors such as age, gender, marital status, living situation, and age at first contact with treatment.

Discussion

Overall findings support the use of SMS text messaging as a means to enhance engagement in treatment-seeking individuals with psychosis. All but one study demonstrated that an SMS text messaging intervention was associated with improved clinical engagement. Moreover, all studies found that no significant harm was associated with SMS text messaging interventions. Similarly, feasibility findings suggest overall patient endorsement for the use of SMS text messaging in the treatment of psychosis.

Medication

Using SMS text messaging for medication reminders appears to have a significant positive effect on both attitudes about medication [47] and medication adherence [47,50,54], with the potential to have lasting effects [49]. However, there are many caveats to these findings. First, findings, although positive and significant, had moderate effect sizes based on small samples, with some trials reporting nonsignificant changes or significant changes for only the subgroups of the study sample. Second, most studies did not investigate effects after the termination of the intervention, highlighting a need for additional longitudinal follow-up studies.

It is also important to acknowledge that medication reminders only address involuntary nonadherence (ie, motivation, forgetfulness, and understanding) and do not directly address voluntary nonadherence (ie, lack of willingness) [57] except potentially indirectly through participants' attitude toward medication. As a cautionary note, this should be considered when evaluating medication reminders as different causes of nonadherence may require different approaches. Furthermore, results indicate SMS text messaging reminders to be most efficacious for patients in a midrange of symptom severity and functioning [47,52,54].

The study that did not find SMS text messaging reminders to have a significant impact on medication adherence used hospital admission rates as the primary outcome [46]. Despite reductions in hospital admissions being one goal of treatment, decreases in hospitalization rates as an indicator of improved medication adherence may be too far removed from the adherence process and may not be an appropriate indicator of the efficacy of SMS text messaging reminders.

The results shown here are similar to those across other areas of health care such that many show positive trends toward improved treatment adherence, with mixed results with regard to significance. For example, some studies find a positive nonsignificant trend toward improved medication adherence in postsurgery populations [58] and individuals with diabetes [33]. Other studies report strong findings in areas such as HIV antiretroviral treatment [32] and general health populations [59]. In contrast, some studies have found no effect of SMS text messaging reminders on populations such as patients with tuberculosis [60].

Some of the aforementioned studies [33,58] as well as many in this review used one-way messaging, which may explain some of the nonsignificant findings. A meta-analysis examining one-way versus two-way SMS text messages as medication reminders found that two-way messaging can increase oral medication adherence in a breadth of medical conditions by 23%, whereas one-way messaging shows little or no effect [61]. In this review, 6 interventions used one-way SMS text messaging reminders, yet only 2 studies found nonsignificant changes in either attendance or medication adherence; thus, one-way reminders may still be effective in psychosis populations.

Attendance

Regarding clinic attendance, SMS text messaging reminders were also found to increase the likelihood of attending appointments [51,52]. The target population in the study by Thomas et al [51] was patients who had not yet had contact with psychosis services, yet the use of SMS text messaging reminders increased the likelihood of clinic attendance two-fold at initial appointments. Similar findings are reported for initial attendance at treatment start in substance use literature [62]. This is notable considering the high rates of disengagement and dropout for patients in early stages of accessing care [14] and, therefore, represent a population at elevated risk of not receiving treatment.

Findings in this review replicate previous positive findings on attendance within various health conditions and settings, including primary care, mental health care, and dental care, where reminders improve attendance and decrease the probability of missed appointments [63-65]. Taken together, the positive effects of SMS text messaging on clinic attendance may have the potential for annual national cost-savings in the millions [66], and large health insurance providers have begun to consider reimbursement for use of digital aids [31].

Therapeutic Rapport

Therapeutic rapport may also be enriched by the addition of SMS text messaging as patients in 1 included study reported improved therapeutic rapport with SMS text messaging interventionists compared with their clinical teams [42]. This may be due to factors relating to perceived availability of the mobile interventionist compared with the clinical team [42], as the mobile interventionist is able to reply to questions and problems in real time and communicate with patients daily. This is an important finding given that therapeutic rapport is important for sustained engagement [14,15].

Feasibility

Feasibility outcomes were assessed by both RCTs and qualitative studies, providing important insights for future implementation. All studies examining user outcomes found SMS text messaging to be feasible, safe, and acceptable, gaining the majority endorsement in each sample [42,44,54-56]. These findings help assuage popular concerns around the feasibility of SMS text messaging interventions as cell phone ownership [55,56] and response rates were high [42,54], most were familiar with SMS text messaging or found it easy to learn and use [49,52,56], and only 1 patient reported increased paranoia around usage of mobile phones [56].

In comparison with other mobile strategies such as mobile apps, SMS text messaging studies reported slightly higher response/usage rates (83%-87%) than mobile app studies (69%-86%) in this population [29], although attrition rates were similar [29]. In a direct comparison of SMS text messaging and mobile apps for daily symptom monitoring, authors noted that the set up used for apps is a more involved process and may affect participation rate; however, the user platform in apps was preferred over SMS text messaging by participants [67].

A clear advantage of SMS text messaging over mobile apps is cell phone versus smartphone ownership, and concerns around the exclusion of economically disadvantaged populations, such as those reliant on government financial support. Cell phone ownership steadily decreases with income and is significantly lower among individuals with severe mental illness (SMI). Ownership among earners of an annual income under US $30,000 per year is 84% and 81% for those with psychosis, yet ownership among the lowest earners drops to 50% and 35%, respectively [29]. Cell phone ownership among SMI populations has been reported at 93% in the United States, yet only 50% to 60% owned smartphones [68,69], and 67.9% in Israel [70]. In developing and newly developed countries such as India, however, smartphone ownership is still only present in a minority (22.7%) of mental health populations [71]. Importantly, trends in smartphone ownership, although still below cell phone ownership rates as of 2018, have been steadily increasing for

years in the general population [72] as well as populations with psychosis [29] and will likely become more ubiquitous as younger populations grow in this digital age.

Strengths and Limitations

A strength of this review is that we took a comprehensive perspective to engagement and provided an inclusive review of clinical engagement in populations with psychosis [14]. However, an even more comprehensive examination of technologically aided engagement could have been created by including other platforms that use mobile devices, such as mobile apps, as a helpful comparison group. This would help further elucidate patient preferences and effectiveness to ascertain information on which platform is best suited for implementation in the treatment of psychosis. Another limitation of this review may be the design itself, as a review is limited to data extracted from the published articles of the included studies, without attempts to contact corresponding authors of the included studies to obtain further information. In addition, this review was designed and registered as a qualitative review rather than a meta-analysis, although we acknowledge that a meta-analysis may have provided some more quantitative insights into the use of SMS text messaging to boost engagement in the treatment of psychosis; the meta-analysis itself would be limited given the small number of published studies and the significant heterogeneity in methods, primary outcomes, and durations of interventions..

Limited conclusions can be drawn from included studies owing to the scarcity of effectiveness trials and novelty of digital strategies in the treatment of psychosis. Additionally, many current trials are pilot studies that are time-limited with small sample sizes using convenience sampling and, in some cases, with underpowered analyses. Another limitation is that no study examined impacts on clinician workflow or cost-effectiveness or privacy considerations beyond educating participants on privacy risks. Furthermore, there are no standard measurements for clinical engagement or any of its domains, making results too heterogeneous to be accurately compared.

Future Directions

The central theme drawn from the presented studies is that individuals with psychosis prefer a person-centered approach, with a personal feel, and that this approach tends to be more effective. Suggestions for future trials using SMS text messaging include the involvement of participants in co-design, employing two-way messaging to engage participants in a conversation and allowing for personal choice and autonomy within the intervention, namely personal choice over content, timing, and frequency. This is especially important given the considerable heterogeneity in psychosis spectrum disorders and the flexibility of SMS text messaging to afford opportunities for individualized treatment approaches. Moreover, digital tools have the potential to increase patient autonomy and independence by extending care outside the clinic and putting it literally in patients' hands.

Given the heterogeneity of clinical presentation and treatment, it should also be considered that not all patients with psychosis are prescribed medication, so standalone medication reminders would not be universally helpful. Other subgroups requiring additional planning include individuals with physical disabilities, low literacy levels, severe cognitive impairments, developmental delays, experiencing homelessness, or share a cell phone with others [42]. Another key area of consideration is safety protocols. Despite a study's report that several urgent issues had to be escalated to the primary care team resulting in expedited prescription fills, home visits, and safety assessments [43], none of the included studies reference precautions or procedures relating to safety. Such issues underscore the importance of developing safety protocols for digital tools.

Finally, due to the potential increase in patient autonomy and independence, SMS text messaging–augmented care could lead to a decreased dependence on clinical staff, allowing clinicians more flexibility to manage caseloads and potentially leading to positive financial and economic outcomes while simultaneously allowing for better, more personalized care. Future studies should, therefore, report on the cost-effectiveness of the interventions as well as impact on workflow to assess how readily SMS text messaging can be implemented.

Conclusions

In sum, this review demonstrates the potential of SMS text messaging to be used as an adjunct platform to support clinical treatment as a means of improved engagement. These findings show that SMS text messaging is extremely well tolerated, safe, and accepted among individuals with psychosis, testifying to its potential to not only improve engagement in care but also to extend care beyond the clinic. To this end, SMS text messaging stands to offer a pragmatic solution to boost clinical engagement and provide an alternative avenue to access treatment. To this end, a more thorough examination of practical considerations, consistent measures for engagement factors, and rigorous examinations of its outcomes (eg, medication blood plasma levels or electronic medication caps compared with self-reporting) are required to adequately assess its efficacy, cost-effectiveness, and readiness for implementation.

Acknowledgments

All author credits have been listed. The authors would also like to acknowledge Heather Cunningham from Gerstein Library Research Services at the University of Toronto who was instrumental in guiding the search strategy and protocol development.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of results for included studies.
[DOCX File , 22 KB - mental_v7i4e16993_app1.docx]

References

1. World Health Organization. World Health Report 2001 - Mental Health: New Understanding, New Hope. Geneva: World Health Organization; 2001.
2. Murali V, Oyebode F. Poverty, social inequality and mental health. Adv Psychiatr Treat 2004;10(3):216-224. [doi: 10.1192/apt.10.3.216]
3. Schoenbaum M, Sutherland J, Chappel A, Azrin S, Goldstein AB, Rupp A, et al. Twelve-month health care use and mortality in commercially insured young people with incident psychosis in the United States. Schizophr Bull 2017 Oct 21;43(6):1262-1272 [FREE Full text] [doi: 10.1093/schbul/sbx009] [Medline: 28398566]
4. Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. J Clin Psychiatry 2006;67(Suppl 9):3-8; discussion 36 [FREE Full text] [doi: 10.4088/jcp.1006e12] [Medline: 16965182]
5. Iyer S, Mustafa S, Gariépy G, Shah J, Joober R, Lepage M, et al. A NEET distinction: youths not in employment, education or training follow different pathways to illness and care in psychosis. Soc Psychiatry Psychiatr Epidemiol 2018 Dec;53(12):1401-1411 [FREE Full text] [doi: 10.1007/s00127-018-1565-3] [Medline: 30094632]
6. Chong HY, Teoh SL, Wu DB, Kotirum S, Chiou CF, Chaiyakunapruk N. Global economic burden of schizophrenia: a systematic review. Neuropsychiatr Dis Treat 2016;12:357-373 [FREE Full text] [doi: 10.2147/NDT.S96649] [Medline: 26937191]
7. Gaebel W, Weinmann S, Sartorius N, Rutz W, McIntyre JS. Schizophrenia practice guidelines: international survey and comparison. Br J Psychiatry 2005 Sep;187:248-255. [doi: 10.1192/bjp.187.3.248] [Medline: 16135862]
8. Pringsheim T, Addington D. Canadian schizophrenia guidelines: introduction and guideline development process. Can J Psychiatry 2017 Sep;62(9):586-593 [FREE Full text] [doi: 10.1177/0706743717719897] [Medline: 28789558]
9. Lehman AF, Lieberman JA, Dixon L, McGlashan TH, Miller AL, Perkins DO, American Psychiatric Association, Steering Committee on Practice Guidelines. Practice guideline for the treatment of patients with schizophrenia, second edition. Am J Psychiatry 2004 Feb;161(2 Suppl):1-56. [Medline: 15000267]
10. Addington D, Abidi S, Garcia-Ortega I, Honer WG, Ismail Z. Canadian guidelines for the assessment and diagnosis of patients with schizophrenia spectrum and other psychotic disorders. Can J Psychiatry 2017 Sep;62(9):594-603 [FREE Full text] [doi: 10.1177/0706743717719899] [Medline: 28730847]
11. Bowden CL, Hirschfeld RM, Keck PE, Gitlin MJ, Suppes T, Thase ME, et al. Psychiatry Online. 2010. Practice Guideline for the Treatment of Patients With Bipolar Disorder Second Edition URL: https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/bipolar.pdf [accessed 2020-02-04]
12. Nosé M, Barbui C, Tansella M. How often do patients with psychosis fail to adhere to treatment programmes? A systematic review. Psychol Med 2003 Oct;33(7):1149-1160. [doi: 10.1017/s0033291703008328] [Medline: 14580069]
13. Killaspy H, Banerjee S, King M, Lloyd M. Prospective controlled study of psychiatric out-patient non-attendance. Characteristics and outcome. Br J Psychiatry 2000 Feb;176:160-165. [doi: 10.1192/bjp.176.2.160] [Medline: 10755054]
14. Doyle R, Turner N, Fanning F, Brennan D, Renwick L, Lawlor E, et al. First-episode psychosis and disengagement from treatment: a systematic review. Psychiatr Serv 2014 May 1;65(5):603-611. [doi: 10.1176/appi.ps.201200570] [Medline: 24535333]
15. Becker KD, Buckingham SL, Rith-Najarian L, Kline ER. The Common Elements of treatment engagement for clinically high-risk youth and youth with first-episode psychosis. Early Interv Psychiatry 2016 Dec;10(6):455-467. [doi: 10.1111/eip.12283] [Medline: 26486257]
16. Tindall R, Francey S, Hamilton B. Factors influencing engagement with case managers: perspectives of young people with a diagnosis of first episode psychosis. Int J Ment Health Nurs 2015 Aug;24(4):295-303. [doi: 10.1111/inm.12133] [Medline: 25976922]
17. Kreyenbuhl J, Nossel IR, Dixon LB. Disengagement from mental health treatment among individuals with schizophrenia and strategies for facilitating connections to care: a review of the literature. Schizophr Bull 2009 Jul;35(4):696-703 [FREE Full text] [doi: 10.1093/schbul/sbp046] [Medline: 19491314]
18. Conus P, Lambert M, Cotton S, Bonsack C, McGorry PD, Schimmelmann BG. Rate and predictors of service disengagement in an epidemiological first-episode psychosis cohort. Schizophr Res 2010 May;118(1-3):256-263. [doi: 10.1016/j.schres.2010.01.032] [Medline: 20206475]
19. Ascher-Svanum H, Zhu B, Faries DE, Salkever D, Slade EP, Peng X, et al. The cost of relapse and the predictors of relapse in the treatment of schizophrenia. BMC Psychiatry 2010 Jan 7;10:2 [FREE Full text] [doi: 10.1186/1471-244X-10-2] [Medline: 20059765]
20. 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 1;175(5):443-452. [doi: 10.1176/appi.ajp.2017.17050480] [Medline: 29495897]

21. Iyer S, Jordan G, MacDonald K, Joober R, Malla A. Early intervention for psychosis: a Canadian perspective. J Nerv Ment Dis 2015 May;203(5):356-364. [doi: 10.1097/NMD.0000000000000288] [Medline: 25900548]
22. Boydell KM, Pong R, Volpe T, Tilleczek K, Wilson E, Lemieux S. Family perspectives on pathways to mental health care for children and youth in rural communities. J Rural Health 2006;22(2):182-188. [doi: 10.1111/j.1748-0361.2006.00029.x] [Medline: 16606432]
23. Ronis ST, Slaunwhite AK, Malcom KE. Comparing strategies for providing child and youth mental health care services in Canada, the United States, and the Netherlands. Adm Policy Ment Health 2017 Nov;44(6):955-966. [doi: 10.1007/s10488-017-0808-z] [Medline: 28612298]
24. Lal S, Malla A. Service engagement in first-episode psychosis: Current issues and future directions. Can J Psychiatry 2015 Aug;60(8):341-345 [FREE Full text] [doi: 10.1177/070674371506000802] [Medline: 26454555]
25. Malla A, Iyer S, McGorry P, Cannon M, Coughlan H, Singh S, et al. From early intervention in psychosis to youth mental health reform: a review of the evolution and transformation of mental health services for young people. Soc Psychiatry Psychiatr Epidemiol 2016 Mar;51(3):319-326. [doi: 10.1007/s00127-015-1165-4] [Medline: 26687237]
26. Nolin M, Malla A, Tibbo P, Norman R, Abdel-Baki A. Early intervention for psychosis in Canada: What is the State of Affairs? Can J Psychiatry 2016 Mar;61(3):186-194 [FREE Full text] [doi: 10.1177/0706743716632516] [Medline: 27254094]
27. Jones KR, Lekhak N, Kaewluang N. Using mobile phones and short message service to deliver self-management interventions for chronic conditions: a meta-review. Worldviews Evid Based Nurs 2014 Apr;11(2):81-88. [doi: 10.1111/wvn.12030] [Medline: 24597522]
28. Firth J, Torous J. Smartphone apps for schizophrenia: a systematic review. JMIR Mhealth Uhealth 2015 Nov 6;3(4):e102 [FREE Full text] [doi: 10.2196/mhealth.4930] [Medline: 26546039]
29. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of 'mHealth' among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar;42(2):448-455 [FREE Full text] [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
30. Ben-Zeev D, Kaiser SM, Brenner CJ, Begale M, Duffecy J, Mohr DC. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. Psychiatr Rehabil J 2013 Dec;36(4):289-296 [FREE Full text] [doi: 10.1037/prj0000019] [Medline: 24015913]
31. Carras MC, Mojtabai R, Furr-Holden CD, Eaton W, Cullen BA. Use of mobile phones, computers and internet among clients of an inner-city community psychiatric clinic. J Psychiatr Pract 2014 Mar;20(2):94-103 [FREE Full text] [doi: 10.1097/01.pra.0000445244.08307.84] [Medline: 24638044]
32. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov 27;376(9755):1838-1845. [doi: 10.1016/S0140-6736(10)61997-6] [Medline: 21071074]
33. Sugita H, Shinohara R, Yokomichi H, Suzuki K, Yamagata Z. Effect of text messages to improve health literacy on medication adherence in patients with type 2 diabetes mellitus: A randomized controlled pilot trial. Nagoya J Med Sci 2017 Aug;79(3):313-321 [FREE Full text] [doi: 10.18999/nagjms.79.3.313] [Medline: 28878436]
34. Lari H, Noroozi A, Tahmasebi R. Comparison of multimedia and SMS education on the physical activity of diabetic patients: an application of health promotion model. Iran Red Crescent Med J 2018;20(s1):e59800. [doi: 10.5812/ircmj.59800]
35. Kaushal T, Montgomery KA, Simon R, Lord K, Dougherty J, Katz LEL, et al. MyDiaText: feasibililty and functionality of a text messaging system for youth with Type 1 Diabetes. Diabetes Educ 2019 Jun;45(3):253-259. [doi: 10.1177/0145721719837895] [Medline: 30902038]
36. Islam SM, Chow CK, Redfern J, Kok C, Rådholm K, Stepien S, et al. Effect of text messaging on depression in patients with coronary heart disease: a substudy analysis from the TEXT ME randomised controlled trial. BMJ Open 2019 Feb 20;9(2):e022637 [FREE Full text] [doi: 10.1136/bmjopen-2018-022637] [Medline: 30787075]
37. Keating SR, McCurry M. Text messaging as an intervention for weight loss in emerging adults. J Am Assoc Nurse Pract 2019 Sep;31(9):527-536. [doi: 10.1097/JXX.0000000000000176] [Medline: 30908408]
38. Song T, Qian S, Yu P. Mobile health interventions for self-control of unhealthy alcohol use: systematic review. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e10899 [FREE Full text] [doi: 10.2196/10899] [Medline: 30694200]
39. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]
40. Kauppi K, Välimäki M, Hätönen HM, Kuosmanen LM, Warwick-Smith K, Adams CE. Information and communication technology based prompting for treatment compliance for people with serious mental illness. Cochrane Database Syst Rev 2014 Jun 17(6):CD009960. [doi: 10.1002/14651858.CD009960.pub2] [Medline: 24934254]
41. Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Joanna Briggs Institute. 2017. JBI Reviewer's Manual - Chapter 3: Systematic reviews of effectiveness URL: https://reviewersmanual.joannabriggs.org/ [accessed 2020-02-04]
42. Ben-Zeev D, Kaiser SM, Krzos I. Remote 'hovering' with individuals with psychotic disorders and substance use: feasibility, engagement, and therapeutic alliance with a text-messaging mobile interventionist. J Dual Diagn 2014;10(4):197-203 [FREE Full text] [doi: 10.1080/15504263.2014.962336] [Medline: 25391277]

43. Aschbrenner KA, Naslund JA, Gill LE, Bartels SJ, Ben-Zeev D. A qualitative study of client-clinician text exchanges in a mobile health intervention for individuals with psychotic disorders and substance use. J Dual Diagn 2016;12(1):63-71 [FREE Full text] [doi: 10.1080/15504263.2016.1145312] [Medline: 26829356]
44. Kauppi K, Kannisto KA, Hätönen H, Anttila M, Löyttyniemi E, Adams CE, et al. Mobile phone text message reminders: measuring preferences of people with antipsychotic medication. Schizophr Res 2015 Oct;168(1-2):514-522. [doi: 10.1016/j.schres.2015.07.044] [Medline: 26293215]
45. Kannisto KA, Adams CE, Koivunen M, Katajisto J, Välimäki M. Feedback on SMS reminders to encourage adherence among patients taking antipsychotic medication: a cross-sectional survey nested within a randomised trial. BMJ Open 2015 Nov 9;5(11):e008574 [FREE Full text] [doi: 10.1136/bmjopen-2015-008574] [Medline: 26553830]
46. Välimäki M, Kannisto KA, Vahlberg T, Hätönen H, Adams CE. Short text messages to encourage adherence to medication and follow-up for people with psychosis (mobile.net): Randomized controlled trial in Finland. J Med Internet Res 2017 Jul 12;19(7):e245 [FREE Full text] [doi: 10.2196/jmir.7028] [Medline: 28701292]
47. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. Psychiatry Res 2012 Dec 30;200(2-3):89-95. [doi: 10.1016/j.psychres.2012.07.034] [Medline: 22901437]
48. Xu D. University Libraries - University of Washington. 2017. Effectiveness of Mobile Text Messaging to Improve Schizophrenia Adherence, Symptoms and Functioning in a Resource-poor Community Setting: 'LEAN' Randomized Controlled Trial URL: https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/40805/Xu_washington_0250E_18081.pdf?sequence=1&isAllowed=y [accessed 2020-02-04]
49. Menon V, Selvakumar N, Kattimani S, Andrade C. Therapeutic effects of mobile-based text message reminders for medication adherence in bipolar I disorder: Are they maintained after intervention cessation? J Psychiatr Res 2018 Sep;104:163-168. [doi: 10.1016/j.jpsychires.2018.07.013] [Medline: 30081390]
50. Beebe L, Smith KD, Phillips C. A comparison of telephone and texting interventions for persons with schizophrenia spectrum disorders. Issues Ment Health Nurs 2014 May;35(5):323-329. [doi: 10.3109/01612840.2013.863412] [Medline: 24766166]
51. Thomas IF, Lawani AO, James BO. Effect of short message service reminders on clinic attendance among outpatients with psychosis at a psychiatric hospital in Nigeria. Psychiatr Serv 2017 Jan 1;68(1):75-80. [doi: 10.1176/appi.ps.201500514] [Medline: 27582239]
52. Pijnenborg GH, Withaar FK, Brouwer WH, Timmerman ME, van den Bosch RJ, Evans JJ. The efficacy of SMS text messages to compensate for the effects of cognitive impairments in schizophrenia. Br J Clin Psychol 2010 Jun;49(Pt 2):259-274. [doi: 10.1348/014466509X467828] [Medline: 19735607]
53. Kravariti E, Reeve-Mates C, da Gama Pires R, Tsakanikos E, Hayes D, Renshaw S, et al. Effectiveness of automated appointment reminders in psychosis community services: a randomised controlled trial. BJPsych Open 2018 Jan;4(1):15-17 [FREE Full text] [doi: 10.1192/bjo.2017.7] [Medline: 29388909]
54. Granholm E, Ben-Zeev D, Link PC, Bradshaw KR, Holden JL. Mobile Assessment and Treatment for Schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. Schizophr Bull 2012 May;38(3):414-425 [FREE Full text] [doi: 10.1093/schbul/sbr155] [Medline: 22080492]
55. Lal S, Dell'Elce J, Tucci N, Fuhrer R, Tamblyn R, Malla A. Preferences of young adults with first-episode psychosis for receiving specialized mental health services using technology: a survey study. JMIR Ment Health 2015;2(2):e18 [FREE Full text] [doi: 10.2196/mental.4400] [Medline: 26543922]
56. Bogart K, Wong SK, Lewis C, Akenzua A, Hayes D, Prountzos A, et al. Mobile phone text message reminders of antipsychotic medication: is it time and who should receive them? A cross-sectional trust-wide survey of psychiatric inpatients. BMC Psychiatry 2014 Jan 22;14:15 [FREE Full text] [doi: 10.1186/1471-244X-14-15] [Medline: 24447428]
57. Steinkamp JM, Goldblatt N, Borodovsky JT, LaVertu A, Kronish IM, Marsch LA, et al. Technological interventions for medication adherence in adult mental health and substance use disorders: a systematic review. JMIR Ment Health 2019 Mar 12;6(3):e12493 [FREE Full text] [doi: 10.2196/12493] [Medline: 30860493]
58. Brix LD, Bjørnholdt KT, Thillemann TM, Nikolajsen L. The effect of text messaging on medication adherence after outpatient knee arthroscopy: a randomized controlled trial. J Perianesth Nurs 2019 Aug;34(4):710-716. [doi: 10.1016/j.jopan.2018.11.011] [Medline: 30852173]
59. Moore DJ, Poquette A, Casaletto KB, Gouaux B, Montoya JL, Posada C, HIV Neurobehavioral Research Program (HNRP) Group. Individualized texting for adherence building (iTAB): improving antiretroviral dose timing among HIV-infected persons with co-occurring bipolar disorder. AIDS Behav 2015 Mar;19(3):459-471 [FREE Full text] [doi: 10.1007/s10461-014-0971-0] [Medline: 25504449]
60. Liu X, Lewis JJ, Zhang H, Lu W, Zhang S, Zheng G, et al. Effectiveness of electronic reminders to improve medication adherence in tuberculosis patients: a cluster-randomised trial. PLoS Med 2015 Sep;12(9):e1001876 [FREE Full text] [doi: 10.1371/journal.pmed.1001876] [Medline: 26372470]
61. Wald DS, Butt S, Bestwick JP. One-way versus two-way text messaging on improving medication adherence: meta-analysis of randomized trials. Am J Med 2015 Oct;128(10):1139.e1-1139.e5. [doi: 10.1016/j.amjmed.2015.05.035] [Medline: 26087045]

62. Kmiec J, Suffoletto B. Implementations of a text-message intervention to increase linkage from the emergency department to outpatient treatment for substance use disorders. J Subst Abuse Treat 2019 May;100:39-44. [doi: 10.1016/j.jsat.2019.02.005] [Medline: 30898326]
63. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014 Oct 17;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]
64. Robotham D, Satkunanathan S, Reynolds J, Stahl D, Wykes T. Using digital notifications to improve attendance in clinic: systematic review and meta-analysis. BMJ Open 2016 Oct 24;6(10):e012116 [FREE Full text] [doi: 10.1136/bmjopen-2016-012116] [Medline: 27798006]
65. Schwebel FJ, Larimer ME. Using text message reminders in health care services: a narrative literature review. Internet Interv 2018 Sep;13:82-104 [FREE Full text] [doi: 10.1016/j.invent.2018.06.002] [Medline: 30206523]
66. Sims H, Sanghara H, Hayes D, Wandiembe S, Finch M, Jakobsen H, et al. Text message reminders of appointments: a pilot intervention at four community mental health clinics in London. Psychiatr Serv 2012 Feb 1;63(2):161-168. [doi: 10.1176/appi.ps.201100211] [Medline: 22302334]
67. Ainsworth J, Palmier-Claus JE, Machin M, Barrowclough C, Dunn G, Rogers A, et al. A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: native smartphone application vs text-messaging only implementations. J Med Internet Res 2013 Apr 5;15(4):e60 [FREE Full text] [doi: 10.2196/jmir.2328] [Medline: 23563184]
68. Gitlow L, Abdelaal F, Etienne A, Hensley J, Krukowski E, Toner M. Exploring the current usage and preferences for everyday technology among people with serious mental illnesses. Occup Ther Ment Heal 2017;33(1):1-14. [doi: 10.1080/0164212X.2016.1211061]
69. Naslund JA, Aschbrenner KA, Bartels SJ. How people with serious mental illness use smartphones, mobile apps, and social media. Psychiatr Rehabil J 2016 Dec;39(4):364-367 [FREE Full text] [doi: 10.1037/prj0000207] [Medline: 27845533]
70. Rahal ZA, Vadas L, Manor I, Bloch B, Avital A. Use of information and communication technologies among individuals with and without serious mental illness. Psychiatry Res 2018 Aug;266:160-167. [doi: 10.1016/j.psychres.2018.05.026] [Medline: 29864616]
71. Sreejith G, Menon V. Mobile phones as a medium of mental health care service delivery: perspectives and barriers among patients with severe mental illness. Indian J Psychol Med 2019;41(5):428-433 [FREE Full text] [doi: 10.4103/IJPSYM.IJPSYM_333_18] [Medline: 31548765]
72. Pew Research Centre. 2019 Jun 12. Mobile Fact Sheet: Mobile Phone Ownership Over Time URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2019-08-10]

Abbreviations

RCT: randomized controlled trial
SMI: Severe Mental Illness

Edited by G Strudwick; submitted 11.11.19; peer-reviewed by N Zaharakis, J Zhang; comments to author 06.12.19; revised version received 03.01.20; accepted 14.01.20; published 02.04.20.

Please cite as:
D'Arcey J, Collaton J, Kozloff N, Voineskos AN, Kidd SA, Foussias G
The Use of Text Messaging to Improve Clinical Engagement for Individuals With Psychosis: Systematic Review
JMIR Ment Health 2020;7(4):e16993
URL: https://mental.jmir.org/2020/4/e16993
doi:10.2196/16993
PMID:32238334

Original Paper

Digital Peer Support Mental Health Interventions for People With a Lived Experience of a Serious Mental Illness: Systematic Review

Karen L Fortuna[1], MSW, PhD; John A Naslund[2], PhD; Jessica M LaCroix[3], PhD; Cynthia L Bianco[4], MS; Jessica M Brooks[5], PhD; Yaara Zisman-Ilani[6], PhD; Anjana Muralidharan[3], PhD; Patricia Deegan[7], PhD

[1]Dartmouth College, Lebanon, NH, United States

[2]Department of Global Health and Social Medicine, Harvard Medical School, Harvard University, Boston, MA, United States

[3]Department of Medical & Clinical Psychology, Uniformed Services University of the Health Sciences, Rockville, MD, United States

[4]The Giesel School of Medicine, Dartmouth College, Concord, NH, United States

[5]Geriatric Research, Education and Clinical Center, James J Peters VA Medical Center, New York, NY, United States

[6]Department of Social and Behavioral Sciences, College of Public Health, Temple University, Philadelphia, PA, United States

[7]Pat Deegan and Associates, Boston, MA, United States

Corresponding Author:
Karen L Fortuna, MSW, PhD
Dartmouth College
46 Centerra Parkway, Suite 200
Lebanon, NH, 03766
United States
Phone: 1 6036533430
Email: klfortuna@gmail.com

Abstract

Background: Peer support is recognized globally as an essential recovery service for people with mental health conditions. With the influx of digital mental health services changing the way mental health care is delivered, peer supporters are increasingly using technology to deliver peer support. In light of these technological advances, there is a need to review and synthesize the emergent evidence for peer-supported digital health interventions for adults with mental health conditions.

Objective: The aim of this study was to identify and review the evidence of digital peer support interventions for people with a lived experience of a serious mental illness.

Methods: This systematic review was conducted using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) procedures. The PubMed, Embase, Web of Science, Cochrane Central, CINAHL, and PsycINFO databases were searched for peer-reviewed articles published between 1946 and December 2018 that examined digital peer support interventions for people with a lived experience of a serious mental illness. Additional articles were found by searching the reference lists from the 27 articles that met the inclusion criteria and a Google Scholar search in June 2019. Participants, interventions, comparisons, outcomes, and study design (PICOS) criteria were used to assess study eligibility. Two authors independently screened titles and abstracts, and reviewed all full-text articles meeting the inclusion criteria. Discrepancies were discussed and resolved. All included studies were assessed for methodological quality using the Methodological Quality Rating Scale.

Results: A total of 30 studies (11 randomized controlled trials, 2 quasiexperimental, 15 pre-post designs, and 2 qualitative studies) were included that reported on 24 interventions. Most of the studies demonstrated feasibility, acceptability, and preliminary effectiveness of peer-to-peer networks, peer-delivered interventions supported with technology, and use of asynchronous and synchronous technologies.

Conclusions: Digital peer support interventions appear to be feasible and acceptable, with strong potential for clinical effectiveness. However, the field is in the early stages of development and requires well-powered efficacy and clinical effectiveness trials.

Trial Registration: PROSPERO CRD42020139037; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=139037

(JMIR Ment Health 2020;7(4):e16460) doi:10.2196/16460

KEYWORDS

peer support; digital mental health; recovery

Introduction

Background

Peer support is recognized globally as an essential recovery service for people with mental health conditions [1]. Peer support services are recovery and wellness support services provided by an individual with a lived experience of recovery from a mental health condition [2]. Peer support is broadly defined as "giving and receiving help founded on key principles of respect, shared responsibility, and mutual agreement of what is helpful" [3], and such services have proven to be instrumental in augmenting traditional mental health treatment [3], thereby providing effective recovery services to people with mental health conditions [4,5]. In particular, peer support services have contributed to increases in patient engagement, positive medical outcomes, patient activation, and greater use of self-management techniques [4,5]. In the largest randomized controlled trial of a peer-led, self-management intervention conducted to date, the researchers found improved physical health and mental health-related quality of life among individuals with serious mental illness and comorbid medical conditions [6]. With the influx of digital mental health services changing the way mental health care is delivered, peer supporters are increasingly using technology to deliver peer support [7].

Digital Peer Support Mental Health Interventions

Traditionally, peer support has been provided as an in-person intervention in multiple service settings such as inpatient and outpatient psychiatric units [3]. More recently, peer support is increasingly being offered through digital technologies, known as digital peer support. Digital peer support is defined as live or automated peer support services delivered through technology media such as peer-to-peer networks on social media, peer-delivered interventions supported by smartphone apps, and asynchronous and synchronous technologies; asynchronous technology facilitates communication between peer support specialists and service users without the need for communication to happen in real time [8]. Through these mobile and online technologies, adoption of digital peer support is expanding the reach of peer support services [8], increasing the impact of peer support without additional in-person sessions [9], and engaging service users in digital mental health [10]. Peers are also co-producing empirically supported digital peer support services [11]. For example, peers working in equal partnership with academic researchers developed [11] and tested a smartphone-based medical and psychiatric self-management intervention for people with mental health conditions, which contributed to statistically significant improvements in psychiatric self-management [9,12]. In addition, improvements were observed in self-efficacy for managing chronic health conditions, hope, quality of life, medical self-management skills, and empowerment [9,12]. Given these advances, there is a need to review and synthesize the emerging evidence for digital peer support interventions for adults with mental health conditions.

Our objectives were (1) to expand on prior reviews that focused on peer support services that did *not* include technology [4,5] or that focused on peer support using technology but *only for* people with psychosis [13] and (2) to conduct a systematic literature review to assess the feasibility, acceptability, and potential effectiveness of digital peer support interventions for adults with serious mental illnesses. We examined the effect of interventions on both biomedical and psychosocial outcomes. In addition, we examined the extent to which researchers engaged service users in the development of the identified digital peer support interventions.

Methods

Search Strategy

We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) procedures [14]. Our search strategy protocol was published to the PROSPERO International prospective register of systematic reviews (Registration number: CRD42020139037). To identify early peer-reviewed articles reporting on digital peer support interventions, we included the following available high-quality electronic reference databases beginning in 1946 until December 2018: PubMed, Embase, Web of Science, Cochrane Central, CINAHL (Cumulative Index to Nursing and Allied Health), and PsycINFO. Each search term was entered as a keyword and assigned the corresponding Medical Subject Heading term (see Multimedia Appendix 1 for the full list of search terms). To identify articles not included in our original search, we reviewed the reference lists of published studies that met the inclusion criteria along with prior systematic reviews, and in June 2019, we searched Google Scholar using different combinations of the search terms.

Study Selection Criteria

Studies were evaluated by the first two authors (KF and JN) who independently screened titles and abstracts. We piloted our title and abstract review protocol on 15 references to ensure 100% concordance/agreement between reviewers before reviewing the entire set of titles and abstracts. These authors independently reviewed all full-text articles meeting the inclusion criteria. Any discrepancies were discussed and resolved. According to the PRISMA guidelines [14], we used the participants, interventions, comparisons, outcomes, and study design (PICOS) criteria [15] to assess study eligibility:

- Participants: Individuals aged≥18 years with either a diagnosis of schizophrenia spectrum disorder (schizophrenia or schizoaffective disorder) or bipolar disorder.
- Intervention: Digital peer support interventions, including peer-delivered interventions, peer augmented interventions, and peer-to-peer social media interventions.
- Comparisons: Studies did not need to have a comparison condition. Interventions could have been delivered at any location such as participants' homes, primary care setting, federally qualified health centers, outpatient facilities,

inpatient facilities, community mental health centers, community settings, or could have been delivered via remote or mobile technology.
- Outcomes: The primary outcomes of interest included those related to feasibility, acceptability, and effectiveness (ie, biomedical and psychosocial outcomes).
- Study design: We included randomized controlled trials, pre-post designs with an experimental or a quasi-experimental comparison condition, qualitative studies, and secondary data analyses if outcomes were relevant to the feasibility, acceptability, and effectiveness of digital peer support interventions. Research protocols, letters to the editor, review articles, pharmacological studies, theoretical articles, and articles that were not peer-reviewed were excluded from this systematic review.

Data Extraction

Relevant data from included studies were extracted in duplicate by two reviewers (KF and JN) using a standardized data collection tool. Prior to data extraction, the two reviewers piloted the data collection tool on five included articles to identify and reconcile inconsistent findings or unintended omission of data. A third reviewer (JB) approved the final set of data, decided on any of the remaining data discrepancies, and extracted study characteristics. Extracted study characteristics included study design, sample size and attrition, participant sociodemographic and clinical characteristics, length of study, description of comparison or control group, physical location of intervention (eg, community mental health centers, Veterans Affairs), a description of the intervention, and outcomes.

In addition to the characteristics listed above, we extracted information regarding the extent to which service users were engaged and participated in the development of intervention components. As no benchmark of participant engagement has been consistently defined in the scientific literature, participation rates were divided to present the spread of data. Participation rates were categorized as high engagement (75% or more engaged throughout the intervention), medium engagement (74% to 50% engage throughout the intervention), and low engagement (49% or less engage throughout the intervention). In the event that percentages were not reported or could not be determined, the authors classified studies based on the information provided (eg, study reported statistically significant levels of engagement).

Studies were further categorized by service delivery type, including peer-to-peer networks, peer-delivered interventions supported by technology, and synchronous and asynchronous technologies.

Methodological Quality Assessment

All included studies were assessed for methodological quality using the Methodological Quality Rating Scale (MQRS) [16], which assesses 12 methodological attributes of quality and has been used in other systematic reviews [17-19]. Cumulative scores range from 0 (poor quality) to 17 (high quality); studies that receive a cumulative score of at least 14 are considered to be high-quality studies [16]. Two authors (JB and CB) independently completed the MQRS for studies that met the inclusion criteria. Discrepancies in MQRS ratings were addressed and resolved by the first two authors (KF and JN).

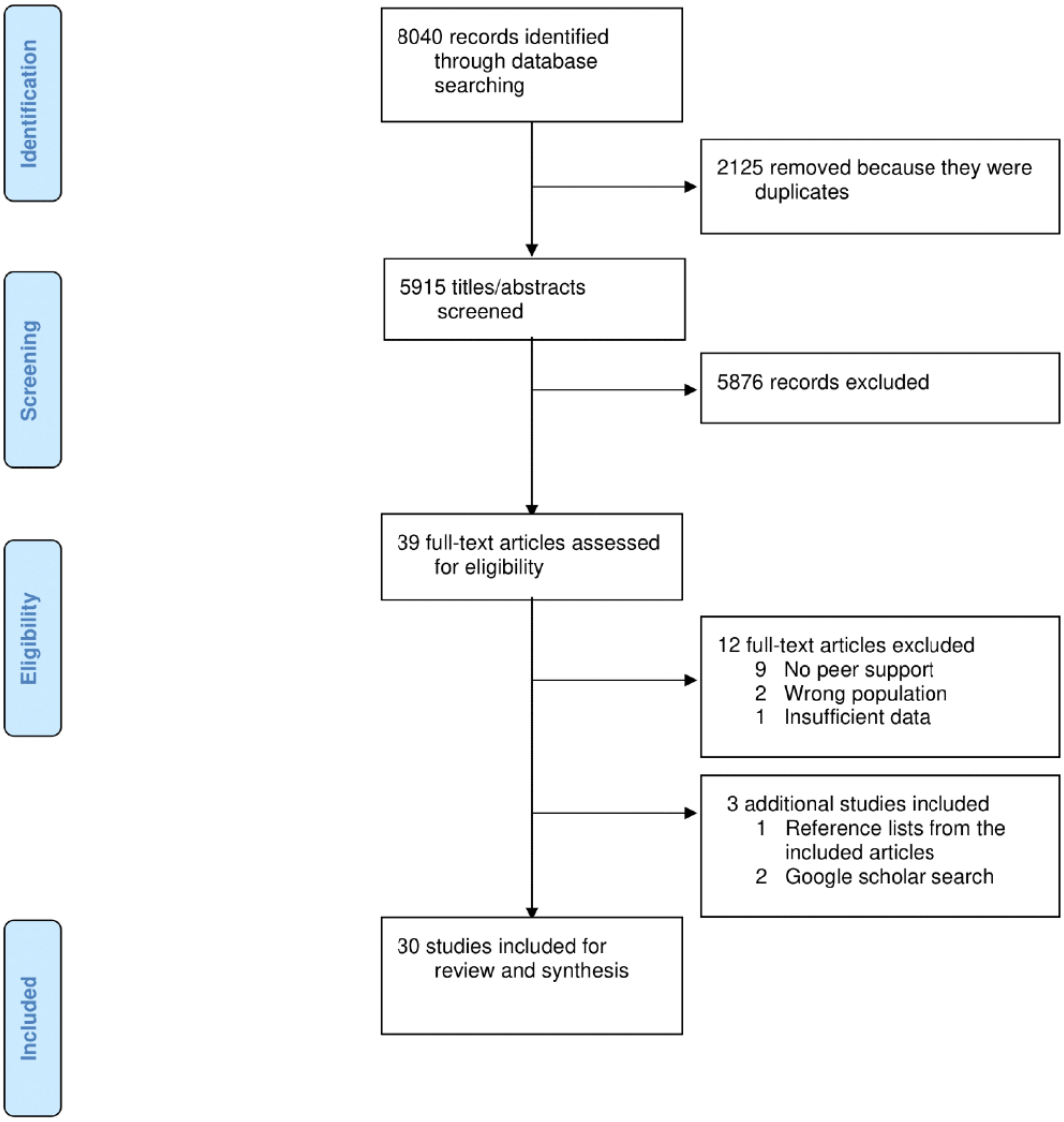
Results

Included Studies

The search strategy identified 8030 articles, including 2125 duplicates. Of the total 5915 titles and abstracts reviewed, 5876 did not meet the inclusion criteria. The full texts of the remaining 39 articles were assessed, and 12 did not meet the inclusion criteria. None of the non–English language articles met the inclusion criteria. Additional articles were found by searching the reference lists of the 27 articles that met the inclusion criteria and conducting a Google Scholar search in June 2019, resulting in an additional 3 included articles. Overall, 30 articles describing 24 interventions met the inclusion criteria and were included in this review (see Figure 1).

As indicated above, included interventions were categorized by the service delivery type by one author (KF). Overall, 14 studies examined peer-to-peer networks, 11 studies examined peer-delivered interventions supported with technology, 2 studies examined peer-supported interventions using synchronous technology, and 3 studies examined peer-supported interventions using asynchronous technology.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of studies included in the review.



Peer-to-Peer Networks

Informal peer support, also known as a "peer-to-peer network" or commonly referred to in the medical community as a "patient-facilitated network," is defined as support given between people with similar life experiences [8]. For example, informal peer support can naturally occur among people in a one-on-one discussion, in a group, or digitally. Informal peer support does not require education or training; rather, people with similar lived experiences define these interactions (see Multimedia Appendix 2). Below, we describe identified studies that were categorized as either stand-alone peer-to-peer networks or peer-to-peer networks combined with evidence-based practices.

Peer-to-Peer Networks

We found one study, a randomized controlled trial, that implemented a peer-to-peer network using a peer support listserv (unmoderated, unstructured, anonymous) and a peer support bulletin board [20]. Although this study was feasible and acceptable, the researchers found no differences between the experimental and control groups on any of the outcomes of interest, including quality of life, empowerment, social support, or psychiatric symptoms.

Peer-to-Peer Networks Combined With Evidence-Based Practices

We found 13 studies that implemented a peer-to-peer network in combination with evidence-based practices [21-31]. These studies included pre-post studies [25-30] and not fully powered randomized controlled trials [21,31] that were designed to address self-management [31], social cognition training [28], weight management [25,26,29,30], motivational enhancement [24], psychoeducation [22,23], or parenting skills training [21,32]. Peer support was facilitated through Facebook [24-26], Google Docs [33], internet-based bulletin boards [20,27], listserv [20,32], or smartphone apps [24,28]. Three studies combined peer-to-peer networks with fitness trackers to promote self-monitoring of physical activity and exercise [25,30].

Overall, these studies appeared feasible. However, attrition rates varied widely. Among the studies that reported attrition, the attrition rates of in-person studies ranged from 0% to 78% [24-28]. Among studies that reported attrition in the technology portion of the study, the attrition rates remained relatively constant: one study using Facebook reported 24% attrition [30] and studies using internet-based bulletin boards reported 0%-5% attrition [22,27].

Participants in studies of interventions consisting of peer-to-peer networks combined with evidence-based practice interventions reported statistically significant improvements in psychiatric symptoms (ie, fewer positive symptoms [22,23] and fewer depression symptoms [24,27]), self-management and biometric outcomes (ie, self-efficacy, weight loss, decreased body mass index [25], clinically significant improvements in cardiovascular fitness [25,26]), person-reported outcomes (ie, improved patient satisfaction [24,26]), service utilization (ie, decreased hospital admissions and hospital length of stay [29]), knowledge (ie, significant increase in knowledge about schizophrenia [22,23]), parenting (ie, improved skills and satisfaction [21,32]), and psychosocial processes (ie, reduced maladaptive social cognitions [28] and improved motivation [24]).

Peer-Delivered Interventions Supported With Technology

Overview

We found 11 studies that implemented peer-delivered interventions supported with technology [9,34-43]. These studies included 1 qualitative study [40], 5 pre-post studies [9,34,36,42,43], 3 quasiexperimental studies [35,38,39], and 2 randomized controlled trials [37,41] that aimed to address integrated medical and psychiatric self-management [9], shared decision making [34,43], cognitive enhancement therapy [41], physical well-being [38,39], and weight management [42]. Peer-delivered services were delivered through smartphone apps [38,43], in-person and augmented by a smartphone app [9,40], in-person and augmented by text messaging and a fitness tracker [42], or via a web-based platform with a peer [34,41] (see Multimedia Appendix 3).

Overall, these studies seem to also be feasible, with the exception of one study [43], in which the mode of delivery (ie, smartphone app) was deemed not feasible. However, attrition rates varied greatly, ranging from 0% to 77% [9,34,35,38,39,41,42].

There was a wide variety of reported outcomes in the peer-delivered interventions supported with technology. Below, we present the statistically significant outcomes, qualitative outcomes, and null results.

Statistically Significant Outcomes

Participants who completed the peer-delivered interventions supported with technology experienced statistically significant benefits in shared decision–making reports [34,37], health care utilization (ie, improvement in engagement in mental health outpatient services [35] and provider perceptions of consumer involvement [37]), self-management (ie, improved medication adherence [35] and psychiatric self-management [9]), person-reported outcomes (ie, improvement in recovery, self-reported psychiatric symptoms [36], lower medication side effects [37]), and patient experience (ie, better relationship and communication between users and doctors [37]). In addition, in one study, the presence of a peer support specialist was associated with better cognitive performance among participants with a lived experience of a serious mental illness completing computerized neurocognitive remediation training sessions [41].

Qualitative Outcomes

Service users and providers reported finding the app in one study useful for supporting people in recovery via its ability to provide an overview of the intervention and set a treatment agenda, while promoting a connection with peer support specialists [40]. Two studies found technological obstacles to the use of technology defined as frustration with technical malfunctions in the app [40,43].

Null Results

Some studies found modest improvements (not statistically significant) in hope, empowerment, social support, quality of life [9], self-reported physical health status [38], and weight loss [42]. Some studies did not find changes in outcomes as related to walking, self-reported global health quality, mental health quality, health control, mental health control, stages of change for exercise [38], self-reported treatment involvement, hope, self-reported patient activation and autonomy preferences [36], patient activation, patient satisfaction, psychiatric distress, global assessment of functioning, drug-induced extrapyramidal symptoms, medication adherence, and quality of life [37]. One study reported low levels of patient satisfaction with an app [39].

Asynchronous and Synchronous Technologies

Synchronous Technologies

As shown in Multimedia Appendix 4, we found 2 articles that reported on a fully powered randomized controlled trial using synchronous technologies [44,45]. In these studies, peer support was facilitated through the telephone [45] combined with internet-based modules accessible at Veteran Affairs clinic kiosks. Overall, the intervention appeared feasible; however, only 86/276 (31.2%) of enrolled participants completed the intervention [44,45]. Participants who attended at least one session of the intervention, whose weight was in the obese range, and who completed synchronous technology components reported statistically significant benefits [45]. An intent-to-treat analysis of all participants found that the synchronous technology intervention increased physical activity [44].

Asynchronous Technologies

Multimedia Appendix 4 also summarizes the studies related to asynchronous technologies. We found 3 studies that implemented asynchronous technologies [46-48], including an exploratory qualitative study [48], a pre-post study [46], and a fully powered randomized controlled trial [44,47]. Peer support was facilitated through (1) peer-led videos in combination with a website to be used on a tablet by mental health workers to structure discussions about personal recovery [44,46], (2) peer-written emails and peer-led videos on recovery in

combination with a noninteractive online psychoeducation program [47], and (3) an interactive website including videos of people with lived experience of mental illness discussing their recovery [46,48]. These interventions aimed at personal recovery [46] and psychiatric self-management [47,48]. Overall, these interventions were feasible, with reports of 80% to 100% engagement [46,47].

Participants who completed the asynchronous technology interventions reported statistically significant benefits in personal recovery [46]. Qualitative findings showed that participants felt "inspired," "knowing I'm not alone," and "believing recovery is possible" [48]. One study compared two versions of a peer-supported intervention with a nonpeer-supported psychoeducational text-based website: one consisting of an online psychoeducational program augmented by video testimony and advice from peers, and another consisting of that same program supplemented with email-delivered peer coaching and support [47]. No significant differences were found in any measures of psychiatric symptomatology, anxiety, perceived control over the illness, perceived stigma, functioning, patient satisfaction, or health locus of control [47].

Community Engagement and Participation

More than half of the studies (16/30, 53%) included community engagement in intervention development [9,22-24,27,31,34-36,38-40,43,46-48]. Four studies used consultative methods of community engagement in intervention development (ie, advice, video content, and information via focus groups) [27,38,47], four studies used active community engagement (ie, co-design as equal partners between scientists and community members) [31,36,46,48], two studies used active and consultative methods [34,35], three studies used a combination of consultative and user-centered designs (ie, focus groups, task analysis, and usability testing) [22-24], two studies used active and user-centered designs (ie, co-design as equal partners between scientists and community members, task analysis, and usability testing) [9], and one study only used a user-centered design [40]. One study did not incorporate any community engagement in intervention development [43]. The remaining studies did not report any community engagement techniques.

Four studies did not report participant engagement in the intervention [25,29,32,40]. In addition, 14 studies were classified as high engagement [9,21,23,24,26,28,31,34-36,38,46,48,49], 7 studies were classified as medium engagement [26,27,30,43-47], and 5 studies were classified as low engagement [20,22,37,39,41]. Among studies that reported high engagement and an intervention development description [23,24,31,34,35,38,47,48], 3 studies included a real-world effectiveness assessment (not efficacy). Of those, studies with the highest level of engagement employed active methods [31,36] or a combination of active and consultative methods [34,35] (see Multimedia Appendix 5).

Methodological Quality Assessment

Methodological quality was evaluated using an adapted version of the MQRS [16]. MQRS total scores ranged from 2 to 12, with a mean score of 7.5 (SD 2.55) and a median score of 8; six studies had a score10, indicating high methodological quality (see Multimedia Appendix 6). Four studies had a score4, indicating low methodological quality. Many of these studies did not report detailed information about methodology (eg, information about control, follow up, dropout, data analysis). Characteristics associated with methodological quality included use of a manualized intervention design (k=9, 69%), provision of sufficient information for replication (k=11, 85%), and inclusion of baseline characteristics (k=10, 77%).

Discussion

Principal Findings

There is growing evidence that digital peer support interventions can improve the lives of people with serious mental illness. This systematic review identified 30 studies that reported on 24 digital peer support interventions. Most of the studies established support for the feasibility, acceptability, and preliminary effectiveness of the interventions with regard to enhancing participants' functioning, reducing symptoms, and improving program utilization. Peer-delivered and technology-supported interventions demonstrated the most promising evidence for both self-reported biomedical and psychosocial outcomes. Attrition rates varied greatly through all digital peer support platforms. Studies with the highest level of digital health engagement employed active community engagement methods or a combination of active and consultative community engagement methods to develop digital peer support interventions.

The evidence base for digital peer support interventions is predominantly built on single-site trials that included small samples and varying follow-up lengths, which greatly restricted the external validity of these interventions in real-world settings. Digital peer support interventions experience the same issues that are common in the field of digital mental health; thus, well-powered and methodologically rigorous studies are needed to confirm the effect of digital peer support interventions. Billions of dollars are being invested in digital innovation; however, many digital innovations are developed by businesses with profit-making interests, not public health interests. These publicly available digital innovations are marketed without adequate evidence of their effectiveness [50]. Academic and peer support specialists partnering with businesses to rigorously and scientifically appraise digital peer support interventions may lead to the next innovations in peer support as well as digital innovations more broadly.

Peer-delivered and technology-supported interventions demonstrated the most promising evidence for biomedical and psychosocial outcomes. Peer-to-peer networks combined with evidence-based practices predominately included biomedical outcome measures and found positive changes such as reductions in psychiatric symptoms [22,27], maladaptive social cognitions [28], and body weight [25]. In contrast, peer-delivered interventions supported with technology found positive changes in *both* biomedical and psychosocial outcomes such as hope, empowerment, social support, quality of life (no statistically significant improvement) [9,38], recovery [36],

medication adherence [35], psychiatric self-management [9], and neurocognitive remediation (statistically significant improvement) [41]. Although the goals of peer support are not typically the same goals as those of traditional clinical services [51], psychosocial outcomes important to service users such as "hope" may act as an important mechanism of health related to biomedical outcomes [52]. Thus, it may be methodologically appropriate to include both biomedical and psychosocial outcomes in order to advance the field of mental health.

Attrition rates varied greatly through all digital peer support platforms. High rates of attrition before achieving intervention effects is a constant challenge in digital psychiatry [53]. Peer support has been noted as a human factor in digital health engagement that facilitates engagement differently than a clinician-patient relationship [10]. For instance, Dr. Fortuna's model of *reciprocal accountability* indicates that peers promote autonomy, flexible expectations, shared lived experience, and bonding within digital interventions. In contrast, Mohr's model of *supportive accountability* purports that clinicians foster a therapeutic alliance, positive perceptions of providers' expertise, and high expectations that the service user has to justify their action or inaction [54]. A prior review found that the addition of technology-mediated peer support might potentially enhance participant engagement and adherence to mental health interventions [13]. However, the varied attrition rates observed in our review suggest more research needs to be done regarding peer support as a human factor in engagement.

Few of the identified studies employed active participation methods such as community-engaged research; rather, the majority of studies employed less involved, consultative methods such as focus groups or requested feedback, or did not report if the community was involved in digital health development. Studies with methodologically appropriate sample sizes to determine effectiveness (ie, outside of a controlled clinical environment or usability testing facilities) that reported the highest level of engagement used active methods such as co-design of the digital programs with peer support specialists as equal partners [31,36] or a combination of active and consultative methods [34,35] (ie, feedback). Applying more participatory research techniques such as community-based participatory research or the Academic-Peer Support Specialists Partnership [11] may facilitate intervention engagement as co-designed interventions become more relevant to the community's specific needs [8]. Use of co-design and participatory techniques may also improve the generalizability of digital mental health interventions, as many existing programs are often tailored to the needs of individuals who will use them rather than for the broader population of individuals who may benefit [55].

Studies included in this systematic review used a wide range of definitions for peer support. A measurement of fidelity for peer support is in early stages of development [56]; however, to date, a fidelity measure of peer support does not exist, despite a national call for such a measure [57]. As such, it is not known what mechanisms of peer support have a positive or negative impact on biomedical and psychosocial outcomes. Potentially, the variation in peer-delivered interventions supported with technology produces such different results because peer support is currently delivered in widely differing ways. This is potentially in contrast to the other interventions that included asynchronous and synchronous technologies, as the technology was used to guide fidelity. There are multiple models of peer support that are based on different theories, principles, and practices, and they include separate training and statewide Medicaid accreditation procedures. For peer support to become widely recognized as an essential services delivery practice and to ensure quality delivery of peer-based services across diverse settings, a measurement of fidelity is needed.

Limitations

We acknowledge several limitations of this study. First, we recognize the lack of longitudinal outcomes identified in the included studies, which did not allow us to assess the impact of digital peer services over time. Further research is needed to determine how to sustain improvements in health, especially as people with serious mental illness may need community-based support to augment traditional outpatient clinical support and prevent premature intervention attrition. It would also be important to determine whether the addition of peer services can further contribute to sustained outcomes for people with a lived experience of a serious mental illness using digital interventions. Second, we cannot reliably differentiate which specific aspects of peer support or other health intervention components contributed to positive changes in biomedical and psychosocial outcomes. This highlights an important area of future research focused on examining the specific peer intervention components or peer service–delivery strategies that produce the best outcomes.

Conclusion

This is the first study to systematically examine digital peer support interventions for people with serious mental illness. It is feasible for peers to use multiple technology modalities to facilitate the delivery of peer support and other evidence-based practices in health care. Similar to other fields in psychiatry, digital health engagement remains an issue. As peer support is an essential recovery service for people with mental health conditions globally [1], this systematic review found that the science of digital peer support is advancing. Advancement of the field requires additional proof-of-concept studies and an examination of digital peer services delivery strategies in combination with high levels of community engagement, as well as further evidence of intervention effectiveness across high, middle, and low-income countries.

Acknowledgments

KF received funding support from National Institutes of Mental Health (NIMH), Grant K01MH117496. This research was also supported by the VA Rehabilitation Research and Development Service (CDA IK2RX002339; AJ, principal investigator). The funders had no role in the study design; collection, analysis, and interpretation of data; writing of the report; and decision to

submit the paper for publication. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Authors' Contributions

KF led the study conceptualization, conducted the systematic review, analyzed the data, drafted the manuscript, and approved the final submitted draft. KF had full access to all the data in the study and had final responsibility for the decision to submit for publication. JN contributed to the design of the systematic review, data acquisition, analysis, and interpretation of data for the work; critically revised the manuscript; and approved the final submitted draft. JL contributed to the design of the systematic review, data acquisition, analysis, methodological quality review and interpretation of data for the work; critically revised the manuscript; and approved the final submitted draft. CB and JB contributed to the methodological quality review and interpretation of data for the work, critically revised the manuscript, and approved the final submitted draft. YZ contributed to the design of the systematic review, data acquisition, and interpretation of data for the work; critically revised the manuscript; and approved the final submitted draft. AM contributed to the design of the systematic review and study conceptualization, critically revised the manuscript, and approved the final submitted draft. PD contributed to the interpretation of findings, critically revised the manuscript, and approved the final submitted draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PubMed search terms for the systematic review. The search was developed for PubMed and was translated to Embase, Web of Science, Cochrane Central, CINAHL, and PsycInfo.
[DOCX File , 14 KB - mental_v7i4e16460_app1.docx]

Multimedia Appendix 2
Current state of the evidence for peer-to-peer networks.
[DOCX File , 20 KB - mental_v7i4e16460_app2.docx]

Multimedia Appendix 3
Current state of evidence for peer-delivered interventions supported with technology.
[DOCX File , 18 KB - mental_v7i4e16460_app3.docx]

Multimedia Appendix 4
Current state of evidence for asynchronous and synchronous technologies.
[DOCX File , 15 KB - mental_v7i4e16460_app4.docx]

Multimedia Appendix 5
Level of community engagement in intervention development and participant engagement rates.
[DOCX File , 19 KB - mental_v7i4e16460_app5.docx]

Multimedia Appendix 6
Assessment of methodological quality of included studies.
[DOCX File , 15 KB - mental_v7i4e16460_app6.docx]

References

1. WHO. Promoting recovery in mental health and related services. World Health Organization 2017.
2. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
3. Mead S, MacNeil C. Peer support: What makes it unique. Int J Psychosoc Rehab 2006;10(2):29-37.
4. Wexler B, Davidson L, Styron T, Strauss J. 40 Years of Academic Public Psychiatry. In: Jacobs S, Griffiths EEH, editors. Severe and persistent mental illness. Hoboken, NJ: John Wiley & Sons; 2007:1-20.
5. Chinman M, George P, Dougherty RH, Daniels AS, Ghose SS, Swift A, et al. Peer support services for individuals with serious mental illnesses: assessing the evidence. Psychiatr Serv 2014 Apr 01;65(4):429-441. [doi: 10.1176/appi.ps.201300244] [Medline: 24549400]

6. Druss BG, Singh M, von Esenwein SA, Glick GE, Tapscott S, Tucker SJ, et al. Peer-Led Self-Management of General Medical Conditions for Patients With Serious Mental Illnesses: A Randomized Trial. Psychiatr Serv 2018 May 01;69(5):529-535 [FREE Full text] [doi: 10.1176/appi.ps.201700352] [Medline: 29385952]
7. Fortuna KL, Aschbrenner KA, Lohman MC, Brooks J, Salzer M, Walker R, et al. Smartphone Ownership, Use, and Willingness to Use Smartphones to Provide Peer-Delivered Services: Results from a National Online Survey. Psychiatr Q 2018 Dec 28;89(4):947-956 [FREE Full text] [doi: 10.1007/s11126-018-9592-5] [Medline: 30056476]
8. Fortuna KL, Venegas M, Umucu E, Mois G, Walker R, Brooks JM. The Future of Peer Support in Digital Psychiatry: Promise, Progress, and Opportunities. Curr Treat Options Psych 2019 Jun 20;6(3):221-231. [doi: 10.1007/s40501-019-00179-7]
9. Fortuna KL, DiMilia PR, Lohman MC, Bruce ML, Zubritsky CD, Halaby MR, et al. Feasibility, Acceptability, and Preliminary Effectiveness of a Peer-Delivered and Technology Supported Self-Management Intervention for Older Adults with Serious Mental Illness. Psychiatr Q 2018 Jun 26;89(2):293-305 [FREE Full text] [doi: 10.1007/s11126-017-9534-7] [Medline: 28948424]
10. Fortuna KL, Brooks JM, Umucu E, Walker R, Chow PI. Peer Support: a Human Factor to Enhance Engagement in Digital Health Behavior Change Interventions. J Technol Behav Sci 2019 May 29;4(2):152-161. [doi: 10.1007/s41347-019-00105-x]
11. Fortuna K, Barr P, Goldstein C, Walker R, Brewer L, Zagaria A, et al. Application of Community-Engaged Research to Inform the Development and Implementation of a Peer-Delivered Mobile Health Intervention for Adults With Serious Mental Illness. J Participat Med 2019 Mar 19;11(1):e12380. [doi: 10.2196/12380]
12. Fortuna KL, Naslund JA, Aschbrenner KA, Lohman MC, Storm M, Batsis JA, et al. Text message exchanges between older adults with serious mental illness and older certified peer specialists in a smartphone-supported self-management intervention. Psychiatr Rehabil J 2019 Mar;42(1):57-63. [doi: 10.1037/prj0000305] [Medline: 30010355]
13. Biagianti B, Quraishi SH, Schlosser DA. Potential Benefits of Incorporating Peer-to-Peer Interactions Into Digital Interventions for Psychotic Disorders: A Systematic Review. Psychiatr Serv 2018 Apr 01;69(4):377-388 [FREE Full text] [doi: 10.1176/appi.ps.201700283] [Medline: 29241435]
14. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
15. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak 2007 Jun 15;7(1):16 [FREE Full text] [doi: 10.1186/1472-6947-7-16] [Medline: 17573961]
16. Miller WR, Wilbourne PL. Mesa Grande: a methodological analysis of clinical trials of treatments for alcohol use disorders. Addiction 2002 Mar;97(3):265-277. [doi: 10.1046/j.1360-0443.2002.00019.x] [Medline: 11964100]
17. Cabassa LJ, Ezell JM, Lewis-Fernández R. Lifestyle interventions for adults with serious mental illness: a systematic literature review. Psychiatr Serv 2010 Aug;61(8):774-782 [FREE Full text] [doi: 10.1176/ps.2010.61.8.774] [Medline: 20675835]
18. Vaughn MG, Howard MO. Integrated psychosocial and opioid-antagonist treatment for alcohol dependence: A systematic review of controlled evaluations. Soc Work Res 2004 Mar 01;28(1):41-53. [doi: 10.1093/swr/28.1.41]
19. Whiteman KL, Naslund JA, DiNapoli EA, Bruce ML, Bartels SJ. Systematic Review of Integrated General Medical and Psychiatric Self-Management Interventions for Adults With Serious Mental Illness. Psychiatr Serv 2016 Nov 01;67(11):1213-1225 [FREE Full text] [doi: 10.1176/appi.ps.201500521] [Medline: 27301767]
20. Kaplan K, Salzer MS, Solomon P, Brusilovskiy E, Cousounis P. Internet peer support for individuals with psychiatric disabilities: A randomized controlled trial. Soc Sci Med 2011 Jan;72(1):54-62. [doi: 10.1016/j.socscimed.2010.09.037] [Medline: 21112682]
21. Kaplan K, Solomon P, Salzer MS, Brusilovskiy E. Assessing an Internet-based parenting intervention for mothers with a serious mental illness: a randomized controlled trial. Psychiatr Rehabil J 2014 Sep;37(3):222-231. [doi: 10.1037/prj0000080] [Medline: 24978623]
22. Rotondi AJ, Haas GL, Anderson CM, Newhill CE, Spring MB, Ganguli R, et al. A Clinical Trial to Test the Feasibility of a Telehealth Psychoeducational Intervention for Persons With Schizophrenia and Their Families: Intervention and 3-Month Findings. Rehabil Psychol 2005 Nov;50(4):325-336 [FREE Full text] [doi: 10.1037/0090-5550.50.4.325] [Medline: 26321774]
23. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/ps.2010.61.11.1099] [Medline: 21041348]
24. Schlosser DA, Campellone TR, Truong B, Etter K, Vergani S, Komaiko K, et al. Efficacy of PRIME, a Mobile App Intervention Designed to Improve Motivation in Young People With Schizophrenia. Schizophr Bull 2018 Aug 20;44(5):1010-1020 [FREE Full text] [doi: 10.1093/schbul/sby078] [Medline: 29939367]
25. Aschbrenner KA, Naslund JA, Shevenell M, Mueser KT, Bartels SJ. Feasibility of Behavioral Weight Loss Treatment Enhanced with Peer Support and Mobile Health Technology for Individuals with Serious Mental Illness. Psychiatr Q 2016 Sep 13;87(3):401-415 [FREE Full text] [doi: 10.1007/s11126-015-9395-x] [Medline: 26462674]

26. Aschbrenner KA, Naslund JA, Shevenell M, Kinney E, Bartels SJ. A Pilot Study of a Peer-Group Lifestyle Intervention Enhanced With mHealth Technology and Social Media for Adults With Serious Mental Illness. J Nerv Ment Dis 2016 Jun;204(6):483-486 [FREE Full text] [doi: 10.1097/NMD.0000000000000530] [Medline: 27233056]
27. Alvarez-Jimenez M, Bendall S, Lederman R, Wadley G, Chinnery G, Vargas S, et al. On the HORYZON: moderated online social therapy for long-term recovery in first episode psychosis. Schizophr Res 2013 Jan;143(1):143-149. [doi: 10.1016/j.schres.2012.10.009] [Medline: 23146146]
28. Biagianti B, Schlosser D, Nahum M, Woolley J, Vinogradov S. Creating Live Interactions to Mitigate Barriers (CLIMB): A Mobile Intervention to Improve Social Functioning in People With Chronic Psychotic Disorders. JMIR Ment Health 2016 Dec 13;3(4):e52 [FREE Full text] [doi: 10.2196/mental.6671] [Medline: 27965190]
29. Gucci F, Marmo F. A study on the effectiveness of E-Mental Health in the treatment of psychosis: Looking to recovery. Eur Psych 2016 Mar;33:S27-S28. [doi: 10.1016/j.eurpsy.2016.01.846]
30. Naslund JA, Aschbrenner KA, Marsch LA, McHugo GJ, Bartels SJ. Facebook for Supporting a Lifestyle Intervention for People with Major Depressive Disorder, Bipolar Disorder, and Schizophrenia: an Exploratory Study. Psychiatr Q 2018 Mar 4;89(1):81-94 [FREE Full text] [doi: 10.1007/s11126-017-9512-0] [Medline: 28470468]
31. Simon GE, Ludman EJ, Goodale LC, Dykstra DM, Stone E, Cutsogeorge D, et al. An online recovery plan program: can peer coaching increase participation? Psychiatr Serv 2011 Jun;62(6):666-669 [FREE Full text] [doi: 10.1176/ps.62.6.pss6206_0666] [Medline: 21632737]
32. O'Shea A, Kaplan K, Solomon P, Salzer MS. Randomized Controlled Trial of an Internet-Based Educational Intervention for Mothers With Mental Illnesses: An 18-Month Follow-Up. Psychiatr Serv 2019 Aug 01;70(8):732-735. [doi: 10.1176/appi.ps.201800391] [Medline: 31023190]
33. O'Leary K, Bhattacharya A, Munson SA, Wobbrock JO, Pratt. Designing peer support chats for mental health. In: Proceedings of the 2017 ACM Conference on Computed Supported Cooperative Work and Social Computing. 2017 Presented at: ACM Conference on Computer Supported Cooperative Work and Social Computing; February 2017; Portland, Oregon p. 1470-1484. [doi: 10.1145/2998181.2998349]
34. Finnerty M, Austin E, Chen Q, Layman D, Kealey E, Ng-Mak D, et al. Implementation and Use of a Client-Facing Web-Based Shared Decision-Making System (MyCHOIS-CommonGround) in Two Specialty Mental Health Clinics. Community Ment Health J 2019 May 13;55(4):641-650 [FREE Full text] [doi: 10.1007/s10597-018-0341-x] [Medline: 30317442]
35. Finnerty MT, Layman DM, Chen Q, Leckman-Westin E, Bermeo N, Ng-Mak DS, et al. Use of a Web-Based Shared Decision-Making Program: Impact on Ongoing Treatment Engagement and Antipsychotic Adherence. Psychiatr Serv 2018 Dec 01;69(12):1215-1221. [doi: 10.1176/appi.ps.201800130] [Medline: 30286709]
36. Salyers MP, Fukui S, Bonfils KA, Firmin RL, Luther L, Goscha R, et al. Consumer Outcomes After Implementing CommonGround as an Approach to Shared Decision Making. Psychiatr Serv 2017 Mar 01;68(3):299-302 [FREE Full text] [doi: 10.1176/appi.ps.201500468] [Medline: 27903137]
37. Yamaguchi S, Taneda A, Matsunaga A, Sasaki N, Mizuno M, Sawada Y, et al. Efficacy of a Peer-Led, Recovery-Oriented Shared Decision-Making System: A Pilot Randomized Controlled Trial. Psychiatr Serv 2017 Dec 01;68(12):1307-1311. [doi: 10.1176/appi.ps.201600544] [Medline: 28945186]
38. Macias C, Panch T, Hicks YM, Scolnick JS, Weene DL, Öngür D, et al. Using Smartphone Apps to Promote Psychiatric and Physical Well-Being. Psychiatr Q 2015 Dec 31;86(4):505-519. [doi: 10.1007/s11126-015-9337-7] [Medline: 25636496]
39. Mueller NE, Panch T, Macias C, Cohen BM, Ongur D, Baker JT. Using Smartphone Apps to Promote Psychiatric Rehabilitation in a Peer-Led Community Support Program: Pilot Study. JMIR Ment Health 2018 Aug 15;5(3):e10092 [FREE Full text] [doi: 10.2196/10092] [Medline: 30111526]
40. Korsbek L, Tønder ES. Momentum: A smartphone application to support shared decision making for people using mental health services. Psychiatr Rehabil J 2016 Jun;39(2):167-172. [doi: 10.1037/prj0000173] [Medline: 27030907]
41. Sandoval LR, González BL, Stone WS, Guimond S, Rivas CT, Sheynberg D, et al. Effects of peer social interaction on performance during computerized cognitive remediation therapy in patients with early course schizophrenia: A pilot study. Schizophr Res 2019 Jan;203:17-23. [doi: 10.1016/j.schres.2017.08.049] [Medline: 28882686]
42. Aschbrenner KA, Naslund JA, Barre LK, Mueser KT, Kinney A, Bartels SJ. Peer health coaching for overweight and obese individuals with serious mental illness: intervention development and initial feasibility study. Transl Behav Med 2015 Sep 11;5(3):277-284 [FREE Full text] [doi: 10.1007/s13142-015-0313-4] [Medline: 26327933]
43. Gulliver A, Banfield M, Morse AR, Reynolds J, Miller S, Galati C. A Peer-Led Electronic Mental Health Recovery App in a Community-Based Public Mental Health Service: Pilot Trial. JMIR Form Res 2019 Jun 04;3(2):e12550 [FREE Full text] [doi: 10.2196/12550] [Medline: 31165708]
44. Muralidharan A, Niv N, Brown CH, Olmos-Ochoa TT, Fang LJ, Cohen AN, et al. Impact of Online Weight Management With Peer Coaching on Physical Activity Levels of Adults With Serious Mental Illness. Psychiatr Serv 2018 Oct 01;69(10):1062-1068 [FREE Full text] [doi: 10.1176/appi.ps.201700391] [Medline: 30041588]
45. Young AS, Cohen AN, Goldberg R, Hellemann G, Kreyenbuhl J, Niv N, et al. Improving Weight in People with Serious Mental Illness: The Effectiveness of Computerized Services with Peer Coaches. J Gen Intern Med 2017 Apr 7;32(Suppl 1):48-55 [FREE Full text] [doi: 10.1007/s11606-016-3963-0] [Medline: 28271427]

46. Thomas N, Farhall J, Foley F, Leitan ND, Villagonzalo K, Ladd E, et al. Promoting Personal Recovery in People with Persisting Psychotic Disorders: Development and Pilot Study of a Novel Digital Intervention. Front Psychiatry 2016 Dec 23;7:196. [doi: 10.3389/fpsyt.2016.00196] [Medline: 28066271]
47. Proudfoot J, Parker G, Manicavasagar V, Hadzi-Pavlovic D, Whitton A, Nicholas J, et al. Effects of adjunctive peer support on perceptions of illness control and understanding in an online psychoeducation program for bipolar disorder: a randomised controlled trial. J Affect Disord 2012 Dec 15;142(1-3):98-105. [doi: 10.1016/j.jad.2012.04.007] [Medline: 22858215]
48. Williams A, Fossey E, Farhall J, Foley F, Thomas N. Recovery After Psychosis: Qualitative Study of Service User Experiences of Lived Experience Videos on a Recovery-Oriented Website. JMIR Ment Health 2018 May 08;5(2):e37 [FREE Full text] [doi: 10.2196/mental.9934] [Medline: 29739737]
49. O'Leary K, Schueller S, Wobbrock J, Pratt W. "Suddenly, we got to become therapists for each other": Designing peer support chats for mental health. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems.: Association for Computing Machinery; 2018 Presented at: 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, Canada. [doi: 10.1145/3173574.3173905]
50. Chinman M, McInnes DK, Eisen S, Ellison M, Farkas M, Armstrong M, et al. Establishing a Research Agenda for Understanding the Role and Impact of Mental Health Peer Specialists. Psychiatr Serv 2017 Sep 01;68(9):955-957 [FREE Full text] [doi: 10.1176/appi.ps.201700054] [Medline: 28617205]
51. Substance Abuse and Mental Health Services Administration. SAMHSA's Working Definition of Recovery. USA: Department of Health and Human Services URL: https://store.samhsa.gov/system/files/pep12-recdef.pdf [accessed 2020-02-02]
52. Mead S, Hilton D, Curtis L. Peer support: a theoretical perspective. Psychiatr Rehabil J 2001;25(2):134-141. [doi: 10.1037/h0095032] [Medline: 11769979]
53. Compton MT, Shim RS. The Social Determinants Of Mental Health. Washington DC: American Psychiatric Publishing; 2020.
54. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
55. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
56. Brewer L, Fortuna K, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the Future: Achieving Health Equity Through Health Informatics and Digital Health. JMIR Mhealth Uhealth 2020 Jan 14;8(1):e14512 [FREE Full text] [doi: 10.2196/14512] [Medline: 31934874]
57. Chinman M, McCarthy S, Mitchell-Miland C, Daniels K, Youk A, Edelen M. Early stages of development of a peer specialist fidelity measure. Psychiatr Rehabil J 2016 Sep;39(3):256-265 [FREE Full text] [doi: 10.1037/prj0000209] [Medline: 27618462]

Abbreviations

MQRS: Methodological Quality Rating Scale
PICOS: participants, interventions, comparisons, outcomes, and study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by J Torous; submitted 14.10.19; peer-reviewed by S Allan, K Myrick; comments to author 21.11.19; revised version received 27.12.19; accepted 01.01.20; published 03.04.20.

Please cite as:
Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, Muralidharan A, Deegan P
Digital Peer Support Mental Health Interventions for People With a Lived Experience of a Serious Mental Illness: Systematic Review
JMIR Ment Health 2020;7(4):e16460
URL: https://mental.jmir.org/2020/4/e16460
doi:10.2196/16460
PMID:32243256

Original Paper

The Performance of Emotion Classifiers for Children With Parent-Reported Autism: Quantitative Feasibility Study

Haik Kalantarian[1,2], BSc, MS, DPhil; Khaled Jedoui[3], BSc; Kaitlyn Dunlap[1,2], BA, MRes; Jessey Schwartz[1,2], BA; Peter Washington[1,2], BA, MSc; Arman Husic[1,2], BSc; Qandeel Tariq[1,2], BSc, MSci; Michael Ning[1,2], BSc; Aaron Kline[1,2], BSc; Dennis Paul Wall[1,2,4], BSc, MSc, DPhil

[1]Department of Pediatrics, Stanford University, Stanford, CA, United States

[2]Department of Biomedical Data Science, Stanford University, Stanford, CA, United States

[3]Department of Mathematics, Stanford University, Stanford, CA, United States

[4]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States

Corresponding Author:
Dennis Paul Wall, BSc, MSc, DPhil
Department of Pediatrics
Stanford University
450 Serra Mall
Stanford, CA, 94305
United States
Phone: 1 (650) 7232300
Email: dpwall@stanford.edu

Abstract

Background: Autism spectrum disorder (ASD) is a developmental disorder characterized by deficits in social communication and interaction, and restricted and repetitive behaviors and interests. The incidence of ASD has increased in recent years; it is now estimated that approximately 1 in 40 children in the United States are affected. Due in part to increasing prevalence, access to treatment has become constrained. Hope lies in mobile solutions that provide therapy through artificial intelligence (AI) approaches, including facial and emotion detection AI models developed by mainstream cloud providers, available directly to consumers. However, these solutions may not be sufficiently trained for use in pediatric populations.

Objective: Emotion classifiers available off-the-shelf to the general public through Microsoft, Amazon, Google, and Sighthound are well-suited to the pediatric population, and could be used for developing mobile therapies targeting aspects of social communication and interaction, perhaps accelerating innovation in this space. This study aimed to test these classifiers directly with image data from children with parent-reported ASD recruited through crowdsourcing.

Methods: We used a mobile game called *Guess What?* that challenges a child to act out a series of prompts displayed on the screen of the smartphone held on the forehead of his or her care provider. The game is intended to be a fun and engaging way for the child and parent to interact socially, for example, the parent attempting to guess what emotion the child is acting out (eg, surprised, scared, or disgusted). During a 90-second game session, as many as 50 prompts are shown while the child acts, and the video records the actions and expressions of the child. Due in part to the fun nature of the game, it is a viable way to remotely engage pediatric populations, including the autism population through crowdsourcing. We recruited 21 children with ASD to play the game and gathered 2602 emotive frames following their game sessions. These data were used to evaluate the accuracy and performance of four state-of-the-art facial emotion classifiers to develop an understanding of the feasibility of these platforms for pediatric research.

Results: All classifiers performed poorly for every evaluated emotion except happy. None of the classifiers correctly labeled over 60.18% (1566/2602) of the evaluated frames. Moreover, none of the classifiers correctly identified more than 11% (6/51) of the angry frames and 14% (10/69) of the disgust frames.

Conclusions: The findings suggest that commercial emotion classifiers may be insufficiently trained for use in digital approaches to autism treatment and treatment tracking. Secure, privacy-preserving methods to increase labeled training data are needed to boost the models' performance before they can be used in AI-enabled approaches to social therapy of the kind that is common in autism treatments.

(JMIR Ment Health 2020;7(4):e13174) doi:10.2196/13174

KEYWORDS

mobile phone; emotion; autism; digital data; mobile app; mHealth; affect; machine learning; artificial intelligence; digital health

Introduction

Background

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by stereotyped and repetitive behaviors and interests as well as deficits in social interaction and communication [1]. In addition, autistic children struggle with facial affect and may express themselves in ways that do not closely resemble those of their peers [2-4]. The incidence of ASD has increased in recent years; it is now estimated that approximately 1 in 40 children in the United States is affected by this condition [5]. Although autism has no cure, there is strong evidence that suggests early intervention can improve speech and communication skills [6].

Common approaches to autism therapy include applied behavior analysis (ABA) and the early start Denver model (ESDM). In ABA therapy, the intervention is customized by a trained behavioral analyst to specifically suit the learner's skills and deficits [7]. The basis of this program is a series of structured activities that emphasize the development of transferable skills to the real world. Similarly, naturalistic developmental behavioral interventions such as ESDM support the development of core social skills through interactions with a licensed behavioral therapist while emphasizing joint activities and interpersonal exchange [8]. Both treatment types have been shown to be safe and effective, with their greatest impact potential occurring during early intervention at younger ages [9-11].

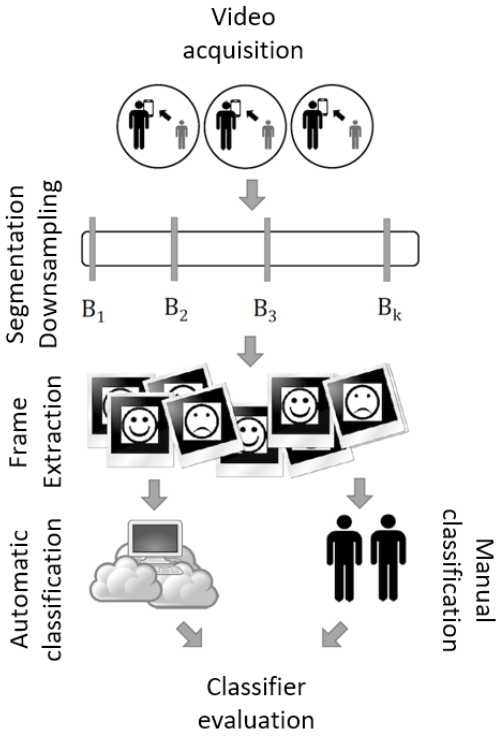
Despite significant progress in understanding this condition in recent years, imbalances in coverage and barriers to diagnosis and treatment remain. In developing countries, studies have noted a lack of trained health professionals, inconsistent treatments, and an unclear pathway from diagnosis to intervention [12-14]. Within the United States, research has shown that children in rural areas receive diagnoses approximately 5 months later than children living in cities [15]. Moreover, it has been observed that children from families near the poverty line receive diagnoses almost a full year later than those from higher-income families. Data-driven approaches have estimated that over 80% of US counties contain no diagnostic autism resources [16]. Even months of delayed access to therapy can limit the effectiveness of subsequent behavioral interventions [15]. Alternative solutions that can ameliorate some of these challenges could be derived from digital and mobile tools. For example, we developed a wearable system using Google Glass that leverages emotion classification algorithms to recognize the facial emotion of a child's conversation partner for real-time feedback and social support and showed treatment efficacy in a randomized clinical trial [17-25].

Various cloud-based emotion classifiers may help the value and reach of mobile tools and solutions. These include four commercially available systems: Microsoft Azure Emotion application programming interface (API) [26], Amazon Rekognition [27], Google Cloud Vision [28], and Sighthound [29]. Whereas most implementations of these emotion recognition APIs are proprietary, these algorithms are typically trained using large facial emotion datasets such as the Cohn-Kanade database [30] and Belfast-Induced Natural Emotion Database [31], which have few examples of children. Due to this bias in labeled examples, it is possible that these models do not generalize well to the pediatric population, including children with developmental delays such as autism, which is evaluated in this study. This study puts the disparity to test. To do so, we use our mobile game *Guess What?* [32-35]. This game (native to Android [36] and iOS [37] platforms) fosters engagement between the child and their social partner, such as a parent, through charades-like games while building a database of facial image data enriched for a range of emotions exhibited by the child during the game sessions.

The primary contributions of this study are as follows:

1. We present a mobile charades game, Guess What?, to crowdsource emotive video from its players. This framework has utility both as a mechanism for the evaluation of existing emotion classifiers and for the development of novel systems that appropriately generalize to the population of interest.
2. We present a study in which 2602 emotive frames are derived from 21 children with a parent-reported diagnosis of autism using data from the Guess What mobile game collected in a variety of heterogeneous environments.
3. The data were used to evaluate the accuracy and performance of several state-of-the-art classifiers using the workflow shown in Figure 1, to develop an understanding of the feasibility of using these APIs in future mobile therapy approaches.

Figure 1. A mobile charades game played between caregiver and child is used to crowdsource emotive video, subsampled and categorized by both manual raters and automatic classifiers. Frames from these videos form the basis of our dataset to evaluate several emotion classifiers.



Related Work

To the best of our knowledge, this is the first work to date that benchmarks public emotion recognition APIs on children with developmental delays. However, a number of interesting apps have been proposed in recent years, which employ vision-based tools or affective computing solutions as an aid for children with autism. The emergence of these approaches motivates a careful investigation of the feasibility of commercial emotion classification algorithms for the pediatric population.

Motivated by the fact that children with autism can experience cognitive or emotional overload, which may compromise their communication skills and learning experience, Picard et al [38] provided an overview of technological advances for sensing autonomic nervous system activation in real-time, including wearable electrodermal activity sensors. A more general overview of the role of affective computing in autism is provided by Kalioby et al [39], with the motivating examples of using technology to help individuals better navigate the socioemotional landscape of their daily lives. Among the enumerated devices include those developed at the Massachusetts Institute of Technology media laboratory, such as *expression glasses* that discriminate between several emotions, skin conductance-sensing gloves for stress detection, and a pressure-sensitive mouse to infer affective state from how individuals interact with the device. Devices made by industry include the SenseWear Pro2 armband, which includes a variety of wearable sensors that can be repurposed for stress and productivity detection, smart gloves that can detect breathing rate and blood pressure, and wireless heart-rate monitors that can be analyzed in the context of environmental stressors [40].

Prior research conducted by us has demonstrated the efficacy of mobile video phenotyping approaches for children with ASD in general [41-47] and via the use of emotion classifiers integrated with the Google Glass platform to provide real-time behavioral support to children with ASD [17-25]. In addition, other studies have confirmed the usability, acceptance, and overall positive impact on families of Google Glass–based systems that use emotion recognition technology to aid social-emotional communication and interaction for autistic children [48,49]. In addition to these efforts, a variety of other smart-glass devices have been proposed. For example, the SenseGlass [50] is among the earliest works that propose leveraging the Google Glass platform to capture and process real-time affective information using a variety of sensors. The authors proposed apps, including the development of affect-based user interfaces, and empowering wearers toward behavioral change through emotion management interventions.

Glass-based affect recognition that predates the Google Glass platform has also been proposed. Scheirer et al [51] used piezoelectric sensors to detect expressions such as confusion and interest, which were detected with an accuracy of 74%. A more recent work proposes a device called *Empathy Glasses* [52] in which users can see, hear, and feel from the perspective of another individual. The system consists of wearable hardware to transmit the wearers' gaze and facial expression and a remote interface where visual feedback is provided, and data are viewed.

The research for smart-glass–based interventions is further supported by other technological systems that have been developed and examined within the context of developmental delays, including the use of augmented reality for object discrimination training [53], assistive robotics for therapy [54-56], and mobile assistive technologies for real-time social

skill learning [57]. Furthermore, the use of computer vision and gamified systems to both detect and teach emotions continues to progress. A computational approach to detect facial expressions optimized for mobile platforms was proposed [58], which demonstrated an accuracy of 95% from a 6-class set of expressions. Leo et al [59] proposed a computational approach to assess the ability of children with ASD to produce facial expressions using computer vision, validated by three expert raters. Their findings demonstrated the feasibility of a human-in-the-loop computer vision system for analyzing facial data from children with ASD. Similar to this study, which utilizes *Guess What?*, a charades-style mobile game to collect emotional face data, Park and colleagues proposed six game design methods for the development of game-driven frameworks in teaching emotions to children with ASD, of which include: observation, understanding, mimicking, and generalization, and supports the use of game play to produce data of value to computer vision approaches for children with autism [60].

Although not all of the aforementioned research studies employ emotion recognition models directly, they are indicative of a general transition from traditional health care practices to modern mobile and digital solutions that leverage recent advances in computer vision, augmented reality, robotics, and artificial intelligence [61]. Thus, the trend motivates our investigation of the efficacy of state-of-the-art vision models on populations with developmental delay.

Methods

Overview

In this section, we describe the architecture of *Guess What?* followed by a description of the methods employed to obtain test data and processing the frames therein to evaluate the performance of several major emotion classifiers. Although dozens of APIs are available, we limit our analysis to some of the most popular systems from major providers of cloud services as a fair representation of the state-of-the-art in publicly available emotion recognition APIs. The systems evaluated in this work were Microsoft Azure Emotion API (Azure) [26], Amazon AWS Rekognition (AWS) [27], Google Cloud Vision API (Google) [28], and Sighthound (SH) [29].

System Architecture

The evaluation of the state-of-the-art in public emotion classification APIs on children with ASD requires a dataset derived from subjects from the relevant population group with a fair amount of consistency in its format and structure. Moreover, as data are limited, it is critical that the video contains a high density of emotive frames to simplify the manual annotation process when establishing a ground truth. Therefore, we have developed and launched an educational mobile game on the Google Play Store [34] and iOS App Store [35], *Guess What?*, from which we derive emotive video.

In this game, parents hold the phone such that the front camera and screen are facing outward toward the child. When the game session begins, the child is shown a prompt that the caregiver must guess based on the child's gestures and facial expressions. After a correct guess is acknowledged, the parent tilts the phone forward, indicating that a point should be awarded. At this time, another prompt is shown. If the one holding the phone cannot make a guess, he/she will tilt the phone backward to skip the frame and automatically proceed to the next. This process repeats until the 90-second game session has elapsed. Meta information is generated for each game session that indicates the times at which various prompts are shown and when the correct guesses occur.

Although a number of varied prompts are available, the two that are most germane to facial affect recognition and emotion recognition are emojis and faces, as shown in Figures 2 and 3, respectively. After the game session is complete, caregivers can elect to share their files and associated metadata to an institution review board-approved secure Amazon S3 bucket that is fully compliant with the Stanford University's high-risk application security standards. A more detailed discussion of the mechanics and applications of *Guess What?* is described in [29-32].

Figure 2. Prompts from the emoji category are caricatures, but many are still associated with the classic Ekman universal emotions.

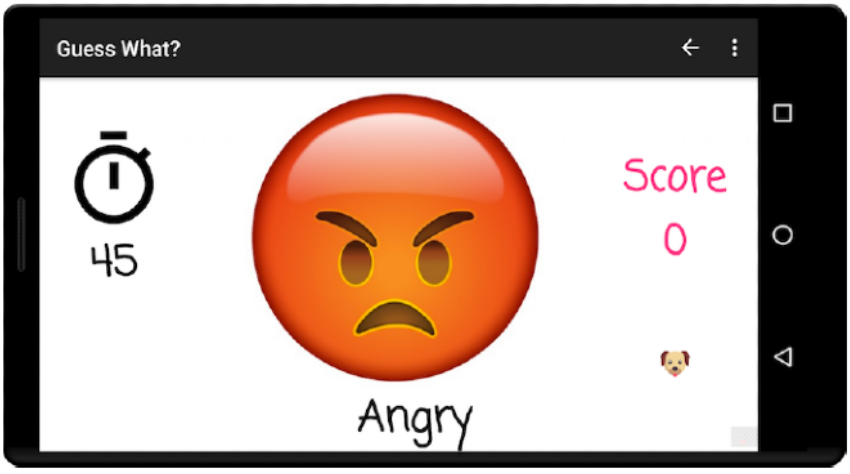
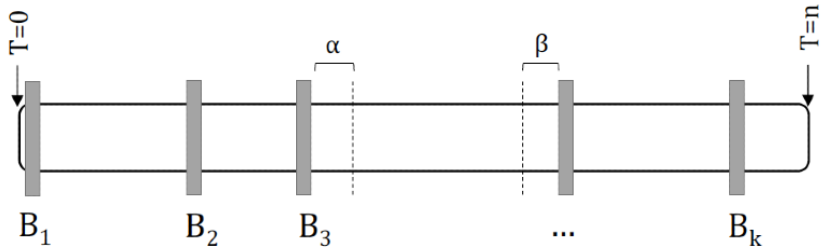


Figure 3. Prompts from the faces category are derived from real photos of children over a solid background.

The structure of a video is shown in Figure 4. Each uploaded video yields *n* video frames, delineated by *k* boundary points, B_1-B_k, where each boundary point represents the time at which a new prompt is shown to the user. To obtain frames associated with a particular emotion, one should first identify the boundary point associated with that emotion through the game meta information, *i*. Having identified this boundary point, frames between B_i and B_i+1 can be associated with this prompt. However, two additional factors remain. It typically takes some time, α, for the child to react after the prompt is shown. Moreover, there is often a time period, β, between the child's acknowledgment of the parents' guess and phone tilt by the parent, during which time the child may adopt a neutral facial expression. Therefore, the frames of interest are those that lie between $B_i+\alpha$ and $B_{i+1}-\beta$.

Figure 4. The structure of a single video is characterized by its boundary points, which identify the times at which various prompts were shown to the child.



The proposed system is centered on two key aims. First, this mechanism facilitates the acquisition of structured emotive videos from children in a manner that challenges their ability to express facial emotion. Whereas other forms of video capture could be employed, a gamified system encourages repeated use and has the potential to contain a much higher density of emotive frames than a typical home video structured around nongaming activities. As manual annotation is employed as a ground truth for evaluating emotion classification, a high concentration of emotive frames within a short time period is essential to the simplification and reduction of the burden associated with this process. A second aim is to potentially facilitate the aggregation of labeled emotive videos from children using a crowdsourcing mechanism. This can be used to augment existing datasets with labeled images or create new ones for the development of novel deep-learning-based emotion classifiers that can potentially overcome the limitations of existing methods.

Data Acquisition

A total of 46 videos from 21 subjects were analyzed in this study. These data were collected over 1 year. Ten videos were collected in a laboratory environment from six subjects with ASD who played several games in a single session administered by a member of the research staff. An additional 36 videos were acquired through crowdsourcing from 15 remote participants. Diagnosis of any form of developmental disorder was provided by the caregiver through self-report during the registration process, along with demographic information (gender, age, ethnicity). The collected information included diagnoses of autistic disorder (autism), ASD, Asperger's syndrome, pervasive developmental disorder (not otherwise specified), childhood disintegrative disorder, no diagnosis, no diagnosis but suspicious, and social communication (pragmatic) disorder. Additionally, a free-text field was available for parents to specify additional conditions. The videos were evaluated by a clinical professional using the Diagnostic and Statistical Manual of Mental Disorders-V criteria before inclusion [1]. Caregivers of all children who participated in the study selected the *autism spectrum disorder* option.

The format of a *Guess What?* gameplay session generally enforces a structure on the derived video: the device is held in landscape mode, the child's face is contained within the frame,

and the distance between the child and camera is typically between 2 and 10 feet. Nevertheless, these videos were carefully screened by members of the research staff to ensure the reliability and quality of the data therein; videos that did not include children, were corrupt, filmed under poor lighting conditions, or did not include plausible demographic information were excluded from the analysis. The average age of the participating children was 7.3 (1.76) years. Due to the small sample size and nonuniform incidence of autism between genders [62], 18 of the 21 participants were male. Although participants explored a variety of game mechanics, all analyzed videos were derived from the two categories most useful for the study of facial affect: faces and emojis. After each game session, the videos were automatically uploaded to an Amazon S3 bucket through an Android background process.

Data Processing

Most emotion classification APIs charge users per an http request, rendering the processing of every frame in a video prohibitive in terms of both time and cost. To simplify our evaluation, we subsampled each video at a rate of two frames per second. These frames formed the basis of our experiments. To obtain ground truth, two raters manually assigned an emotion label to each frame based on the seven Ekman universal emotions [63], with the addition of a neutral label. Some frames were discarded when there was no face, or the quality was too poor to make an assessment. A classifier's performance on a frame was evaluated only under the conditions that the frame was valid (of sufficient quality), and the two manual raters agreed on the emotion label associated with the frame. Frames were considered of insufficient quality if: (1) the frame was too blurry to discern, (2) the child was not in the frame, (3) the image or video was corrupt, or (4) there were multiple individuals within the frame.

From a total of 5418 reviewed frames, 718 were discarded due to a lack of agreement between the manual raters. An additional 2123 frames were discarded because at least one rater assigned the *not applicable (N/A)* label, indicating that the frame was of insufficient quality. This was due to a variety of factors but generally caused by motion artifacts or the child leaving the frame due to the phone being tilted in acknowledgment of a correct guess. The total number of analyzed frames was 2602 divided between the categories shown in Table 1.

As shown, most frames were *neutral*, with a preponderance of *happy* frames in the nonneutral category. Owing to the limited number of *scared* and *confused* frames, this emotion was omitted from our analysis. We also merged the *contempt* and *anger* categories due to their similarity of affect and streamline analysis.

Table 1. The distribution of frames per category (N=2602).

Emotion	Frames, n
Neutral	1393
Emotive	1209
Happy	864
Sad	60
Surprised	165
Disgusted	69
Angry	51

As not all emotion classifiers represented their outputs in a consistent format, some further simplifications were made in our analysis. First, it was necessary to make minor corrections to the format of the outputted data. For example, *happy* and *happiness* were considered identical. In the case of AWS, the *confused* class was ignored, as many other classifiers did not support it. Moreover, *calm* was renamed *neutral*. As AWS, Azure, and Sighthound returned probabilities rather than a single label, a frame in which no emotion class was associated with a probability of over 70% was considered a failure. For Google Vision, classification confidence was associated with a categorical label rather than a percentage. In this case, frames did not receive an emotion classification as *likely* or *very likely* were considered failures. It is also worth noting that this platform, unlike all the others, does not contain *disgust* or *neutral* classes. The final emotions evaluated in this study were *happy*, *sad*, *surprise*, *anger*, *disgust*, and *neutral*, with the latter two omitted for Google Cloud Vision.

As real-time use is an important aspect of mobile therapies and aids, we evaluated the performance of each classifier by calculating the number of seconds required to process each 90-second video subsampled to one frame per second. This evaluation was performed on a Wi-Fi network tested with an average download speed of 51 Mbps and an average upload speed of 62.5 Mbps. For each classifier, this experiment was repeated 10 times to obtain the average amount of time required to process the subsampled video.

Results

Overview

In this section, we present the results of our evaluation of *Guess What?* as well as the performance of the evaluated classifiers: Microsoft Azure Emotion API (Azure) [26], AWS [27], Google Cloud Vision API (Google) [28], and SH [29]. Abbreviations for emotions described within this section can be found in Textbox 1.

Textbox 1. Abbreviations for emotions.

HP: Happy

CF: Confused

N/A: Not applicable

SC: Scared

SP: Surprised

DG: Disgusted

AG: Angry

Classifier Accuracy

Comparison With Ground Truth (Classifiers)

Table 2 shows the performance of each classifier calculated by the percentage of correctly identified frames compared with the ground truth for categories *neutral*, *emotive*, and *all*. A *neutral* frame is one in which the face is recognized, and the neutral label is assigned high confidence. Any other frame within the categories of *happy*, *sad*, *surprised*, *disgusted*, and *angry*, are considered emotive frames. A more detailed breakdown of performance by emotion is shown in Table 3. Note that as before, Google's API does not support the *neutral* and *disgust* categories.

Table 2. Percentage of frames correctly identified by classifier: Azure (Azure Cognitive Services), AWS (Amazon Web Services), SH (Sighthound), and Google (Google Cloud Vision). These results only include frames in which there was a face, and the two manual raters agreed on the class. Google Vision API does not support the neutral label.

Classifier	Frame type		
	Emotive (n=1209), n (%)	Neutral (n=1393), n (%)	All (n=2602), n (%)
Azure	798 (66.00)	744 (53.40)	1542 (59.26)
AWS[a]	829 (68.56)	679 (48.74)	1508 (57.95)
Google	785 (64.92)	N/A[b]	N/A
Sighthound	664 (54.92)	902 (64.75)	1566 (60.18)

[a]AWS: Amazon AWS Rekognition.

[b]N/A: not applicable.

Table 3. Percentage of frames correctly identified by emotion type by each classifier: Azure (Azure Cognitive Services), AWS (Amazon Web Services), SH (Sighthound), and Google (Google Cloud Vision). These results only include frames in which there was a face, and the two manual raters agreed on the class. Note: Google Vision API does not support the neutral or disgust labels.

Classifier	Frame type					
	Neutral (n=1394), n (%)	Happy (n=864), n (%)	Sad (n=60), n (%)	Surprised (n=165), n (%)	Disgusted (n=69), n (%)	Angry (n=51), n (%)
AWS	679 (48.74)	709 (82.0)	19 (31)	94 (56.9)	4 (5)	3 (5)
Sighthound	902 (64.75)	545 (63.0)	13 (21)	90 (54.5)	10 (14)	6 (11)
Azure	744 (53.41)	695 (80.4)	20 (33)	80 (48.4)	0 (0)	3 (5)
Google	N/A[a]	676 (78.2)	10 (16)	93 (56.3)	N/A	6 (11)

[a]N/A: not applicable.

Interrater Reliability (Classifiers)

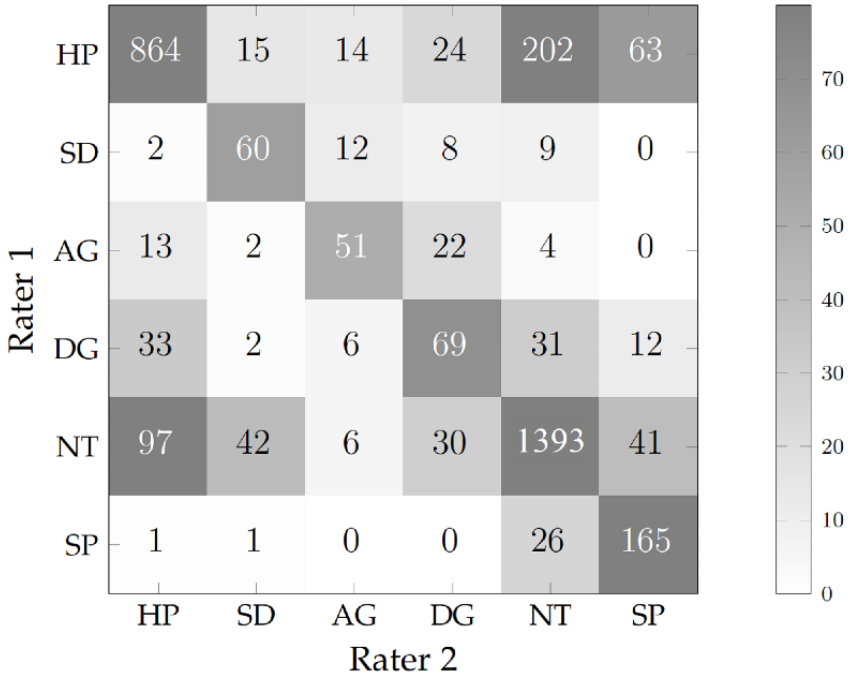
The Cohen kappa statistic is a measure of interrater reliability that factors in the percentage of agreement due to chance; an important consideration when the possible classes are few in number. Figure 5 shows the agreement between every pair of classifiers based on their Cohen kappa score calculated based on every evaluated frame, in which a score of 1 indicates perfect agreement. The results reflect low agreement between most combinations of classifiers. This is particularly true for the lack of agreement between Google and Sighthound, with a Cohen kappa score of 0.2. This is likely because of differences in how the classifiers are tuned for precision and recall; Sighthound correctly identified more *neutral* frames than the others, but performance was lower for the most predominant emotive label: *happy*.

Figure 5. The Cohen's Kappa Score is a measure of agreement between two raters, and was calculated for all four evaluated classifiers: Azure (Azure Cognitive Services), AWS (Amazon Web Services), SH (Sighthound), and Google (Google Cloud Vision). Results indicate weak agreement between all pairs of classifiers.

Interrater Reliability (Human Raters)

The Cohen kappa coefficient for agreement between the two manual raters was 0.74, which was higher than any combination of automatic classifiers evaluated in this study. Although this indicates substantial agreement, it is worth exploring the characteristics of frames in which there was disagreement between the two raters. The full confusion matrix can be seen in Figure 6, which shows the distribution of all frames evaluated by the raters. The results indicate that most discrepancies were between happy and neutral. These discrepancies were likely subtle differences in how the raters perceived a happy face due to the inherent subjectivity of this process. A lack of agreement can also be seen between the disgust-anger categories.

Figure 6. The distribution of frames between the two human raters for each emotion: HP (Happy), SD (Sad), AG (Angry), DG (Disgust), NT (Neutral), and SC (Scared).



Classifier Speed

Wearable and mobile solutions for autism generally require efficient classification performance to provide real-time feedback to users. In some cases, this may be environmental feedback, as in the Autism Glass [17-25], which uses the outward-facing camera of Google Glass to read the emotions of those around the child and provide real-time social queues.

In the case of *Guess What?* the phone's front camera is used to read the expression of the child, which can be analyzed to determine if the facial expression matches the prompt displayed at that time.

To determine if real-time classification performance is feasible with computation offloaded implementations of emotion classifiers, we measured the amount of time required to process a 90-second video recorded at 30 frames per second and subsampled to one frame per second, yielding a total of 90 frames. For each classifier, this experiment was repeated 10 times to obtain the average number of seconds required to process the subsampled video. Table 4 shows the speed of the API-based classifiers used in this study. The values shown in this table represent the amount of time necessary to send each frame to the Web service via an http post request and receive an http response with the emotion label. These frames were processed sequentially, with no overlap between http requests.

Table 4. Speed of the evaluated classifiers.

Classifier	Time (seconds)
Azure	28.6
AWS	90.6
Google	55.9
Sighthound	41.1

The findings indicated that the fastest classifier was Azure, processing all 90 frames in a total of 28.6 seconds. Using Azure with a fast internet connection, it may be possible to obtain semi real-time emotion classification performance, a time of 28.6 seconds corresponds to 3.14 frames per second, which is within the bounds of what could be considered real time. The slowest classifier was AWS, which processed these 90 frames in 90.6 seconds. This corresponds to a frame rate of 0.99 frames per second. In summary, real-time or semi–real-time performance is possible with Web-based emotion classifiers on fast Wi-Fi internet connections. For cellular connections or apps that require frame rates beyond three frames per second, these approaches may be insufficient.

Discussion

Classifier Performance

Results indicate that Google and AWS produced the highest percentage of correctly classified emotive frames, whereas Sighthound produced the highest percentage of correctly identified neutral frames. Google's API did not provide a *neutral* label and, therefore, could not be evaluated. The best system in terms of overall classification accuracy was Sighthound by a small margin, with 60.18% (1566/2602) of the frames correctly identified. Further results indicate that none of the classifiers performed well for any category besides *happy*, which was the emotion most represented in the dataset, as shown in Table 1. In addition, there appears to be a systematic bias toward high recall and low precision for the *happy* category: those classifiers that identified most of the *happy* frames performed worse for those in the *neutral* category.

In summary, the data suggest that although a frame with a smile will be correctly identified in most cases, the ability of the evaluated classifiers to identify other expressions for children with ASD is dubious and presents an obstacle in the design of emotion-based mobile and wearable outcome measures, screening tools, and therapies.

Analysis of Frames

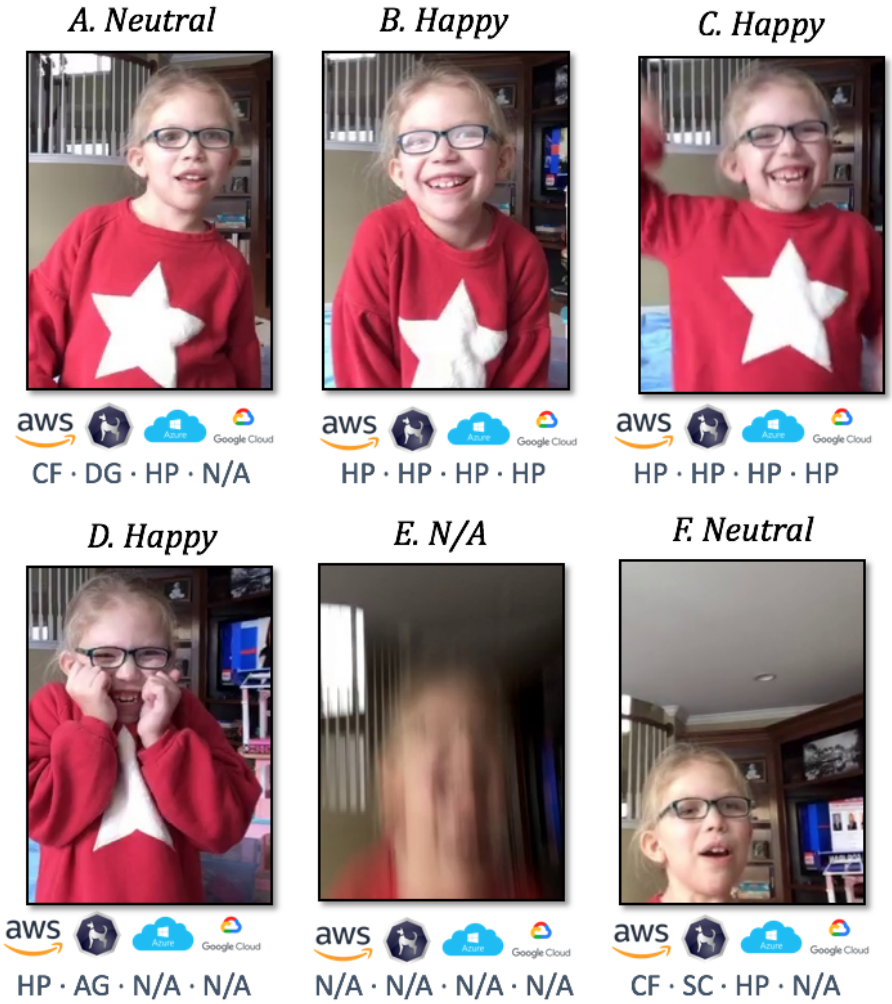
Figure 7 shows six frames from one study participant, reproduced with permission from the child's parents. The top of each frame lists the gold-standard annotation in which both raters agreed on a suitable label for the frame. The bottom of each frame enumerates the labels assigned from each classifier in order: Amazon Rekognition, Sighthound, Azure Cognitive Services, and Google Cloud Vision AI. It should be noted that, as before, these labels are normalized for comparison because each classifier outputs data in a particular format. For example, *N/A* from one classifier could be compared with a blank field in another, whereas some such as Google Cloud explicitly state *Not Sure*; for our purposes, all three of these scenarios were labeled as *N/A* during analysis.

Frame A shows a frame that was labeled as *neutral* by the raters, although each classifier provided a different label: *confused*, *disgusted*, *happy*, and *N/A*. This is an example of a false-positive, detecting an emotion in a neutral frame. A similar example is shown in frame F; most classifiers failed to identify the neutral label. Such false positives are particularly problematic as the *neutral* label is the most prevalent, as shown in Table 1.

In contrast, frames B and C are examples in which the labels assigned by each classifier matched the labels assigned by the manual raters; all classifiers correctly identified the *happy* label. As shown in Table 1, *happy* was the most common nonneutral emotion by a considerable margin, and most classifiers performed quite well in this category; AWS, Azure, and Sighthound all correctly identified between 78.2% (676/864) and 82.0% (709/864) of these frames, although at the expense of increased false-positives such as those shown in frames A and F. An example of a *happy* frame that was identified as such by the human raters but incorrectly by most classifiers is frame D. It is possible that the child's hands covering part of her face contribute to this error, as the frame is otherwise quite similar to frame B. Finally, frame E is an example of a frame that was processed by the classifiers but not included in our experimental results because the human raters flagged the frame as insufficient

due to motion artifacts. In this case, all four classifiers correctly determined that the frame could not be processed.

Figure 7. A comparison of the performance of each classifier on a set of frames highlights scenarios that may lead to discrepancies in the classifier outputs for various emotions: HP (Happy), CF (Confused), DG (Disgust), N/A (Not Applicable), AG (Angry), SC (Scared). Ground truth manual labels are shown on top, with labels derived from each classifier on the bottom.



Limitations

There are several limitations associated with this study, which will be addressed in future work. First, we analyzed only a subset of existing emotion classifiers, emphasizing those from providers of major cloud services. Future efforts will extend this evaluation to include those that are less prolific and require paid licenses. A second limitation is the use of parent-reported diagnoses, which may not always be factual. A third limitation is that although we ruled out some comorbid conditions, we did not rule out all comorbid conditions, including Attention-Deficit/Hyperactivity Disorder, which has been shown to impact emotional processing and function in children [64]. A fourth limitation stems from the lack of neurotypical children. Finally, the dataset we used included an unequal distribution of frames across emotion categories. In the future, we will investigate ways to gather equal numbers of frames, and if this distribution may be related to social deficits associated with autism, or increased prevalence of happy and neutral due to the inherent nature of gameplay. Although our results support the conclusion that the commercial emotion classifiers tested here are not yet at a level needed for use with autistic children, it remains unclear how these models will perform with a larger, more diverse, and stratified sample.

Conclusions

In this feasibility study, we evaluated the performance of four emotion recognition classifiers on children with ASD: Google Cloud Vision, Amazon Rekognition, Microsoft Azure Emotion API, and Sighthound. The average percentage of correctly identified emotive and neutral frames for all classifiers combined was 63.60% (769/1209) and 55.63% (775/1393), respectively, varying greatly between classifiers based on how their sensitivity and specificity were tuned. The results also demonstrated that while most classifiers were able to consistently identify *happy* frames, the performance for *sad*, *disgust*, and *anger* was poor: no classifier identified more than one-third of the frames from either of these categories. We conclude that the performance of the evaluated classifiers is not yet at the level for use in mobile and/or wearable therapy solutions for autistic children, necessitating the development of larger training datasets from these populations to develop more domain-specific models.

Acknowledgments

This work was supported in part by funds to DW from the National Institute of Health (1R01EB025025-01 and 1R21HD091500-01), the Hartwell Foundation, Bill and Melinda Gates Foundation, Coulter Foundation, and program grants from Stanford's Human Centered Artificial Intelligence Program, Precision Health and Integrated Diagnostics Center, Beckman Center, Bio-X Center, Predictives and Diagnostics Accelerator Spectrum, the Wu Tsai Neurosciences Institute Neuroscience: Translate Program, Stanford Spark, and the Weston Havens Foundation. We also acknowledge generous support from Bobby Dekesyer and Peter Sullivan. Dr Haik Kalantarian would like to acknowledge the support from the Thrasher Research Fund and Stanford NLM Clinical Data Science program (T-15LM007033-35).

Conflicts of Interest

DW is the founder of Cognoa. This company is developing digital solutions for pediatric behavioral health, including neurodevelopmental conditions such as autism that are detected and treated using techniques, including emotion classification. AK works as a part-time consultant for Cognoa. All other authors declare no competing interests.

References

1. American Psychological Association. Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition. Arlington, VA: American Psychiatric Pub; 2013.
2. Lozier LM, Vanmeter JW, Marsh AA. Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. Dev Psychopathol 2014 Nov;26(4 Pt 1):933-945. [doi: 10.1017/S0954579414000479] [Medline: 24915526]
3. Loth E, Garrido L, Ahmad J, Watson E, Duff A, Duchaine B. Facial expression recognition as a candidate marker for autism spectrum disorder: how frequent and severe are deficits? Mol Autism 2018;9:7 [FREE Full text] [doi: 10.1186/s13229-018-0187-7] [Medline: 29423133]
4. Fridenson-Hayo S, Berggren S, Lassalle A, Tal S, Pigat D, Bölte S, et al. Basic and complex emotion recognition in children with autism: cross-cultural findings. Mol Autism 2016;7:52 [FREE Full text] [doi: 10.1186/s13229-016-0113-9] [Medline: 28018573]
5. Kogan MD, Vladutiu CJ, Schieve LA, Ghandour RM, Blumberg SJ, Zablotsky B, et al. The prevalence of parent-reported autism spectrum disorder among US children. Pediatrics 2018 Dec;142(6) [FREE Full text] [doi: 10.1542/peds.2017-4161] [Medline: 30478241]
6. Rogers SJ. Brief report: early intervention in autism. J Autism Dev Disord 1996 Apr;26(2):243-246. [doi: 10.1007/bf02172020] [Medline: 8744493]
7. Cooper JO, Heron TE, Heward WL. Applied Behavior Analysis. New Jersey: Pearson; 2007.
8. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics 2010 Jan;125(1):e17-e23 [FREE Full text] [doi: 10.1542/peds.2009-0958] [Medline: 19948568]
9. Dawson G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev Psychopathol 2008;20(3):775-803. [doi: 10.1017/S0954579408000370] [Medline: 18606031]
10. Dawson G, Jones EJ, Merkle K, Venema K, Lowy R, Faja S, et al. Early behavioral intervention is associated with normalized brain activity in young children with autism. J Am Acad Child Adolesc Psychiatry 2012 Nov;51(11):1150-1159 [FREE Full text] [doi: 10.1016/j.jaac.2012.08.018] [Medline: 23101741]
11. Dawson G, Bernier R. A quarter century of progress on the early detection and treatment of autism spectrum disorder. Dev Psychopathol 2013 Nov;25(4 Pt 2):1455-1472. [doi: 10.1017/S0954579413000710] [Medline: 24342850]
12. Scherzer AL, Chhagan M, Kauchali S, Susser E. Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. Dev Med Child Neurol 2012 Dec;54(12):1079-1084 [FREE Full text] [doi: 10.1111/j.1469-8749.2012.04348.x] [Medline: 22803576]
13. van Cong T, Weiss B, Toan KN, Le Thu TT, Trang NT, Hoa NT, et al. Early identification and intervention services for children with autism in Vietnam. Health Psychol Rep 2015;3(3):191-200 [FREE Full text] [doi: 10.5114/hpr.2015.53125] [Medline: 27088123]
14. Samms-Vaughan ME. The status of early identification and early intervention in autism spectrum disorders in lower- and middle-income countries. Int J Speech Lang Pathol 2014 Feb;16(1):30-35. [doi: 10.3109/17549507.2013.866271] [Medline: 24397842]
15. Mandell DS, Novak MM, Zubritsky CD. Factors associated with age of diagnosis among children with autism spectrum disorders. Pediatrics 2005 Dec;116(6):1480-1486 [FREE Full text] [doi: 10.1542/peds.2005-0185] [Medline: 16322174]
16. Ning M, Daniels J, Schwartz J, Dunlap K, Washington P, Kalantarian H, et al. Identification and quantification of gaps in access to autism resources in the United States: an infodemiological study. J Med Internet Res 2019 Jul 10;21(7):e13094 [FREE Full text] [doi: 10.2196/13094] [Medline: 31293243]
17. Washington P, Wall D, Voss C, Kline A, Haber N, Daniels J, et al. SuperpowerGlass: A Wearable Aid for the At-Home Therapy of Children with Autism. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2017 Presented at: ACM'17; June 26-30, 2017; Cambridge, MA. [doi: 10.1145/3130977]

18. Washington P, Catalin V, Nick H, Serena T, Jena D, Carl F, et al. A Wearable Social Interaction Aid for Children with Autism. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI EA'16; May 7-12, 2016; California, USA p. 2348-2354. [doi: 10.1145/2851581.2892282]
19. Daniels J, Schwartz J, Haber N, Voss C, Kline A, Fazel A, et al. 5.13 Design and efficacy of a wearable device for social affective learning in children with autism. J Am Acad Child Psy 2017 Oct;56(10):S257. [doi: 10.1016/j.jaac.2017.09.296]
20. Daniels J, Schwartz JN, Voss C, Haber N, Fazel A, Kline A, et al. Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. NPJ Digit Med 2018;1:32 [FREE Full text] [doi: 10.1038/s41746-018-0035-3] [Medline: 31304314]
21. Daniels J, Haber N, Voss C, Schwartz J, Tamura S, Fazel A, et al. Feasibility testing of a wearable behavioral aid for social learning in children with autism. Appl Clin Inform 2018 Jan;9(1):129-140 [FREE Full text] [doi: 10.1055/s-0038-1626727] [Medline: 29466819]
22. Voss C, Schwartz J, Daniels J, Kline A, Haber N, Washington P, et al. Effect of wearable digital intervention for improving socialization in children with autism spectrum disorder: a randomized clinical trial. JAMA Pediatr 2019 May 1;173(5):446-454 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.0285] [Medline: 30907929]
23. Voss C, Washington P, Haber N, Kline A, Daniels J, Fazel A, et al. Superpower Glass: Delivering Unobtrusive Real-Time Social Cues in Wearable Systems. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016 Presented at: UbiComp'16; September 12-16, 2016; Heidelberg, Germany p. 1218-1226. [doi: 10.1145/2968219.2968310]
24. Voss C, Haber N, Wall DP. The potential for machine learning-based wearables to improve socialization in teenagers and adults with autism spectrum disorder-reply. JAMA Pediatr 2019 Sep 9 [Online ahead of print]. [doi: 10.1001/jamapediatrics.2019.2969] [Medline: 31498377]
25. Kline A, Voss C, Washington P, Haber N, Schwartz H, Tariq Q, et al. Superpower glass. GetMobile 2019 Nov 14;23(2):35-38. [doi: 10.1145/3372300.3372308]
26. Microsoft Azure. Face: An AI Service That Analyzes Faces in Images URL: https://azure.microsoft.com/en-us/services/cognitive-services/face/ [accessed 2020-03-17]
27. Amazon Web Services (AWS) - Cloud Computing Services. Amazon Rekognition URL: https://aws.amazon.com/rekognition/ [accessed 2018-06-17]
28. Google Cloud: Cloud Computing Services. Vision AI URL: https://cloud.google.com/vision/ [accessed 2018-06-17]
29. Sighthound. 2018 Jun 17. Detection API & Recognition API URL: https://www.sighthound.com/products/cloud [accessed 2018-06-17]
30. Cohn J. The Robotics Institute Carnegie Mellon University. 1999. Cohn-Kanade AU-Coded Facial Expression Database URL: https://www.ri.cmu.edu/project/cohn-kanade-au-coded-facial-expression-database/ [accessed 2020-03-24]
31. Sneddon I, McRorie M, McKeown G, Hanratty J. The Belfast induced natural emotion database. IEEE Trans Affective Comput 2012 Jan;3(1):32-41. [doi: 10.1109/t-affc.2011.26]
32. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess what? J Healthc Inform Res 2018 Oct;3(1):43-66. [doi: 10.1007/s41666-018-0034-9]
33. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall D. A Gamified Mobile System for Crowdsourcing Video for Autism Research. In: Proceedings of the 2018 IEEE International Conference on Healthcare Informatics. 2018 Presented at: ICHI'18; June 4-7, 2018; New York, NY, USA p. 350-352. [doi: 10.1109/ichi.2018.00052]
34. Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, et al. Labeling images with facial emotion and the potential for pediatric healthcare. Artif Intell Med 2019 Jul;98:77-86 [FREE Full text] [doi: 10.1016/j.artmed.2019.06.004] [Medline: 31521254]
35. Kalantarian H, Jedoui K, Washington P, Wall DP. A mobile game for automatic emotion-labeling of images. IEEE Trans Games 2018:1 epub ahead of print. [doi: 10.1109/tg.2018.2877325]
36. Google Play. Guess What? URL: https://play.google.com/store/apps/details?id=walllab.guesswhat [accessed 2018-06-01]
37. Apple Store. Guess What? (Wall Lab) URL: https://apps.apple.com/us/app/guess-what-wall-lab/id1426891832 [accessed 2020-03-18]
38. Picard RW. Future affective technology for autism and emotion communication. Philos Trans R Soc Lond B Biol Sci 2009 Dec 12;364(1535):3575-3584 [FREE Full text] [doi: 10.1098/rstb.2009.0143] [Medline: 19884152]
39. el Kaliouby R, Picard R, Baron-Cohen S. Affective computing and autism. Ann N Y Acad Sci 2006 Dec;1093:228-248. [doi: 10.1196/annals.1382.016] [Medline: 17312261]
40. Andre D, Pelletier R, Farringdon J, Er S, Talbott WA, Phillips PP, et al. Semantic Scholar. 2006. The Development of the SenseWear Armband, A Revolutionary Energy Assessment Device to Assess Physical Activity and Lifestyle URL: https://www.semanticscholar.org/paper/The-Development-of-the-SenseWear-%C2%AE-armband-%2C-a-to-Andre-Pelletier/e9e115cb6f381a706687982906d45ed28d40bbac [accessed 2020-03-24]
41. Tariq Q, Daniels J, Schwartz JN, Washington P, Kalantarian H, Wall DP. Mobile detection of autism through machine learning on home video: A development and prospective validation study. PLoS Med 2018 Nov;15(11):e1002705 [FREE Full text] [doi: 10.1371/journal.pmed.1002705] [Medline: 30481180]

42. Tariq Q, Fleming SL, Schwartz JN, Dunlap K, Corbin C, Washington P, et al. Detecting developmental delay and autism through machine learning models using home videos of Bangladeshi children: Development and validation study. J Med Internet Res 2019 Apr 24;21(4):e13822 [FREE Full text] [doi: 10.2196/13822] [Medline: 31017583]
43. Washington P, Kalantarian H, Tariq Q, Schwartz J, Dunlap K, Chrisman B, et al. Validity of online screening for autism: crowdsourcing study comparing paid and unpaid diagnostic tasks. J Med Internet Res 2019 May 23;21(5):e13668 [FREE Full text] [doi: 10.2196/13668] [Medline: 31124463]
44. Washington P, Paskov KM, Kalantarian H, Stockham N, Voss C, Kline A, et al. Feature selection and dimension reduction of social autism data. Pac Symp Biocomput 2020;25:707-718 [FREE Full text] [Medline: 31797640]
45. Abbas H, Garberson F, Glover E, Wall DP. Machine learning approach for early detection of autism by combining questionnaire and home video screening. J Am Med Inform Assoc 2018 Aug 1;25(8):1000-1007. [doi: 10.1093/jamia/ocy039] [Medline: 29741630]
46. Levy S, Duda M, Haber N, Wall DP. Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. Mol Autism 2017;8:65 [FREE Full text] [doi: 10.1186/s13229-017-0180-6] [Medline: 29270283]
47. Fusaro VA, Daniels J, Duda M, DeLuca TF, D'Angelo O, Tamburello J, et al. The potential of accelerating early detection of autism through content analysis of YouTube videos. PLoS One 2014;9(4):e93533 [FREE Full text] [doi: 10.1371/journal.pone.0093533] [Medline: 24740236]
48. Sahin NT, Keshav NU, Salisbury JP, Vahabzadeh A. Second version of Google Glass as a wearable socio-affective aid: positive school desirability, high usability, and theoretical framework in a sample of children with autism. JMIR Hum Factors 2018 Jan 4;5(1):e1 [FREE Full text] [doi: 10.2196/humanfactors.8785] [Medline: 29301738]
49. Keshav NU, Salisbury JP, Vahabzadeh A, Sahin NT. Social communication coaching smartglasses: well tolerated in a diverse sample of children and adults with autism. JMIR Mhealth Uhealth 2017 Sep 21;5(9):e140 [FREE Full text] [doi: 10.2196/mhealth.8534] [Medline: 28935618]
50. Hernandez J, Picard W. SenseGlass: Using Google Glass to Sense Daily Emotions. In: Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology. 2014 Presented at: UIST'14; October 5 - 8, 2014; Hawaii USA p. 77-78. [doi: 10.1145/2658779.2658784]
51. Scheirer J, Fernandez R, Picard R. Expression Glasses: A Wearable Device for Facial Expression Recognition. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 1999 Presented at: CHI'99; May 15-20, 1999; Pittsburgh, Pennsylvania p. 262-263. [doi: 10.1145/632716.632878]
52. Katsutoshi M, Kunze K, Billinghurst M. Empathy Glasses. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI EA'16; May 7-12, 2016; San Jose, California p. 1257-1263. [doi: 10.1145/2851581.2892370]
53. Escobedo L, Tentori M, Quintana E, Favela J, Garcia-Rosas D. Using augmented reality to help children with autism stay focused. IEEE Pervasive Comput 2014 Jan;13(1):38-46. [doi: 10.1109/mprv.2014.19]
54. Scassellati B, Admoni H, Matarić M. Robots for use in autism research. Annu Rev Biomed Eng 2012;14:275-294. [doi: 10.1146/annurev-bioeng-071811-150036] [Medline: 22577778]
55. Feil-Seifer D, Matarić MJ. Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. In: Khatib O, Kumar V, Pappas GJ, editors. Experimental Robotics: The Eleventh International Symposium. Berlin, Heidelberg: Springer; 2009:201-210.
56. Robins B, Dautenhahn K, Boekhorst RT, Billard A. Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Univ Access Inf Soc 2005 Jul 8;4(2):105-120. [doi: 10.1007/s10209-005-0116-3]
57. Escobedo L, Nguyen DH, Boyd LA, Hirano S, Rangel A, Garcia-Rosas D, et al. MOSOCO: A Mobile Assistive Tool to Support Children With Autism Practicing Social Skills in Real-Life Situations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI'12; May 5-10, 2012; Austin, USA p. 2589-2598. [doi: 10.1145/2207676.2208649]
58. Palestra G, Pettiniccho A, Del CM, Carcagni P. Improved Performance in Facial Expression Recognition Using 32 Geometric Features. In: Proceedings of the International Conference on Image Analysis and Processing. 2015 Presented at: ICIAP'15; September 7-11, 2015; Genoa, Italy p. 518-528. [doi: 10.1007/978-3-319-23234-8_48]
59. Leo M, Carcagnì P, Distante C, Spagnolo P, Mazzeo P, Rosato A, et al. Computational assessment of facial expression production in ASD children. Sensors (Basel) 2018 Nov 16;18(11) [FREE Full text] [doi: 10.3390/s18113993] [Medline: 30453518]
60. Park JH, Abirached B, Zhang Y. A Framework for Designing Assistive Technologies for Teaching Children With ASDs Emotions. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 2012 Presented at: CHI EA'12; May 5 - 10, 2012; Texas, Austin, USA p. 2423-2428. [doi: 10.1145/2212776.2223813]
61. Washington P, Park N, Srivastava P, Voss C, Kline A, Varma M, et al. Data-driven diagnostics and the potential of mobile artificial intelligence for digital therapeutic phenotyping in computational psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 2019 Dec 13 [Online ahead of print]. [doi: 10.1016/j.bpsc.2019.11.015] [Medline: 32085921]

62. Christensen DL, Baio J, Braun KN, Bilder D, Charles J, Constantino JN, Centers for Disease Control and Prevention (CDC). Prevalence and characteristics of autism spectrum disorder among children aged 8 years--autism and developmental disabilities monitoring network, 11 sites, United States, 2012. MMWR Surveill Summ 2016 Apr 1;65(3):1-23. [doi: 10.15585/mmwr.ss6503a1] [Medline: 27031587]
63. Ekman P, Friesen WV, O'Sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, et al. Universals and cultural differences in the judgments of facial expressions of emotion. J Pers Soc Psychol 1987 Oct;53(4):712-717. [doi: 10.1037//0022-3514.53.4.712] [Medline: 3681648]
64. Sinzig J, Morsch D, Lehmkuhl G. Do hyperactivity, impulsivity and inattention have an impact on the ability of facial affect recognition in children with autism and ADHD? Eur Child Adolesc Psychiatry 2008 Mar;17(2):63-72. [doi: 10.1007/s00787-007-0637-9] [Medline: 17896119]

Abbreviations

ABA: applied behavior analysis
AI: artificial intelligence
API: application programming interface
ASD: autism spectrum disorder
ESDM: early start Denver model
SH: Sighthound

Edited by G Eysenbach; submitted 18.12.18; peer-reviewed by M Leo, N Sahin; comments to author 27.04.19; revised version received 03.07.19; accepted 23.02.20; published 01.04.20.

Please cite as:
Kalantarian H, Jedoui K, Dunlap K, Schwartz J, Washington P, Husic A, Tariq Q, Ning M, Kline A, Wall DP
The Performance of Emotion Classifiers for Children With Parent-Reported Autism: Quantitative Feasibility Study
JMIR Ment Health 2020;7(4):e13174
URL: https://mental.jmir.org/2020/4/e13174
doi:10.2196/13174
PMID:32234701

Original Paper

Long-Term Effectiveness and Cost-Effectiveness of Videoconference-Delivered Cognitive Behavioral Therapy for Obsessive-Compulsive Disorder, Panic Disorder, and Social Anxiety Disorder in Japan: One-Year Follow-Up of a Single-Arm Trial

Kazuki Matsumoto[1], PhD; Sayo Hamatani[1,2], PhD; Kazue Nagai[3], PhD; Chihiro Sutoh[4], MD, PhD; Akiko Nakagawa[1], MD, PhD; Eiji Shimizu[1,4], MD, PhD

[1]Research Center for Child Mental Development, Chiba University, Chiba, Japan

[2]Japan Society for the Promotion of Science, Tokyo, Japan

[3]Research and Education Center of Health Sciences, Graduate School of Health Sciences, Gunma University, Gunma, Japan

[4]Department of Cognitive Behavioral Physiology, Graduate School of Medicine, Chiba University, Chia, Japan

Corresponding Author:
Kazuki Matsumoto, PhD
Research Center for Child Mental Development
Chiba University
Inohana 1-8-1, Chuo-ku
Chiba,
Japan
Phone: 81 08 043 226 2975
Fax: 81 08 043 226 8588
Email: axpa0219@chiba-u.jp

Abstract

Background: Face-to-face individual cognitive behavioral therapy (CBT) and internet-based CBT (ICBT) without videoconferencing are known to have long-term effectiveness for obsessive-compulsive disorder (OCD), panic disorder (PD), and social anxiety disorder (SAD). However, videoconference-delivered CBT (VCBT) has not been investigated regarding its long-term effectiveness and cost-effectiveness.

Objective: The purpose of this study was to investigate the long-term effectiveness and cost-effectiveness of VCBT for patients with OCD, PD, or SAD in Japan via a 1-year follow-up to our previous 16-week single-arm study.

Methods: Written informed consent was obtained from 25 of 29 eligible patients with OCD, PD, and SAD who had completed VCBT in our clinical trial. Participants were assessed at baseline, end of treatment, and at the follow-up end points of 3, 6, and 12 months. Outcomes were the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS), Panic Disorder Severity Scale (PDSS), Liebowitz Social Anxiety Scale (LSAS), Patient Health Questionnaire–9 (PHQ-9), General Anxiety Disorder–7 (GAD-7), and EuroQol-5D-5L (EQ-5D-5L). To analyze long-term effectiveness, we used mixed-model analysis of variance. To analyze cost-effectiveness, we employed relevant public data and derived data on VCBT implementation costs from Japanese national health insurance data.

Results: Four males and 21 females with an average age of 35.1 (SD 8.6) years participated in the 1-year follow-up study. Principal diagnoses were OCD (n=10), PD (n=7), and SAD (n=8). The change at 12 months on the Y-BOCS was −4.1 (F_1=4.45, P=.04), the change in PDSS was −4.4 (F_1=6.83, P=.001), and the change in LSAS was −30.9 (F_1=6.73, P=.01). The change in the PHQ-9 at 12 months was −2.7 (F_1=7.72, P=.007), and the change in the GAD-7 was −3.0 (F_1=7.09, P=.009). QALY at 12 months was 0.7469 (SE 0.0353, 95% Cl 0.6728-0.821), and the change was a significant increase of 0.0379 (P=.01). Total costs to provide the VCBT were ¥60,800 to ¥81,960 per patient. The set threshold was ¥189,500 ($1723, €1579, and £1354) calculated based on willingness to pay in Japan.

Conclusions: VCBT was a cost-effective way to effectively treat Japanese patients with OCD, PD, or SAD.

Trial Registration: University Hospital Medical Information Network Clinical Trials Registry UMIN000026609; https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000030495

(JMIR Ment Health 2020;7(4):e17157) doi:10.2196/17157

KEYWORDS

long-term effectiveness; cost-effectiveness; videoconference-delivered cognitive behavioral therapy; internet-based cognitive behavioral therapy; obsessive-compulsive disorder; panic disorder; social anxiety disorder

Introduction

Background

Obsessive-compulsive disorder (OCD), panic disorder (PD), and social anxiety disorder (SAD) are mental health illnesses that create severe obstacles for patients in their daily lives [1]. The long-term effectiveness of treatment is worth evaluating, because OCD, PD, and SAD often recur even after improvement following treatment [2-4]. In particular, it is important to guide effective health care policy in countries such as Japan, which have instituted universal public health care insurance systems [5], to optimize limited resources and maintain medical services in consideration of cost-effectiveness.

Telepsychiatry can be delivered to established therapy patients in developed countries where there is wide availability of information and communication devices and internet use is high. Within telepsychiatry, videoconference-delivered cognitive behavioral therapy (VCBT) has proved promising, with the potential to improve the accessibility of specialized care to patients with OCD, PD, and SAD [6]. Even with simple Web cameras, the internet, and information and communication equipment, psychiatrists can significantly improve symptoms by properly examining patients with mental illness, delivering psychological education, and dispensing medication [7]. Multiple clinical trials have reported significant reductions in symptoms of depression, OCD, PD, and SAD as a result of VCBT [6,8,9]. However, we know little about the long-lasting (12 or more months) effectiveness and cost-effectiveness of VCBT, despite its proven short-term effectiveness [4,9,10].

VCBT requires a videoconferencing system, thereby making it more expensive compared with face-to-face cognitive behavioral therapy (CBT). For facilities that provide health care services, VCBT is a little more expensive than traditional CBT. However, for patients, VCBT is less burdensome than face-to-face CBT, as there are no travel costs or time costs associated with hospital visits. VCBT puts the burden of cost on the facility; thus, it is particularly important to assess whether its adoption is a worthwhile approach from the perspective of efficient health care policy.

Objectives of the Study

This study's main objectives were to assess the long-term effectiveness of VCBT for patients with OCD, PD, or SAD and estimate its cost-effectiveness in Japan.

Methods

Study Design

In this study, we included data from our previous clinical trials and follow-ups [6]. We obtained written consent from participants in two stages. First, we obtained participants' written consent forms to research feasibility of VCBT at face-to-face screening before the intervention. Second, those who consented to participate in the follow-up study were requested to resend signed consent forms provided to the researchers. The questionnaires on symptomology were sent by mail or the data collected telephonically at 3, 6, 8, and 12 months after the end of VCBT. These data were used in this study.

In March 2018, the Cognitive Behavioral Therapy Center at Chiba University Hospital implemented a prospective observational study involving all patients who participated in VCBT (reference number: G28038, UMIN000026609) [6]. The study was registered with University Hospital Medical Information Network Clinical Trials Registry [UMIN000026609]. In a follow-up study after the intervention, the institutional review board of Chiba University approved the study protocol (No. 3048).

Participants and Eligibility Criteria in the Clinical Trial

All participants had received face-to-face treatment from the attending physician (psychiatrist) during a previous clinical trial period [6]. VCBT was provided in addition to ongoing face-to-face treatment including pharmacotherapy. Inclusion criteria for our previous clinical trial included informed consent to participate in the study; having a primary diagnosis of OCD, PD, or SAD based on the Mini-International Neuropsychiatric Interview [11,12]; being aged between 19 and 65 years; and having access to the internet at home [6].

Outcomes

Symptomatology

The following Japanese version of three scales were used to assess the severity of the three disorders. The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) was used to measure OCD by identifying the patient' contents of obsessions and compulsions on the symptom checklist and assessing their severity in 4 stages using responses to 10 items on the symptom severity scale [13,14]. The Panic Disorder Severity Scale (PDSS) was used when PD was the principal diagnosis [15,16]. PDSS is a 7-item questionnaire on frequency of panic attacks, extent of subjective distress, impact on daily life, and so on, with response options ranging from 0 to 4 in severity. The Liebowitz Social Anxiety Scale (LSAS) was used for

participants whose principal diagnosis was SAD [17,18]. LSAS is a 24-item questionnaire intended to evaluate the extent of anxiety and avoidance in social situations where social anxiety is noticeable (eg, public speaking, talking to strangers).

We also assessed depression and general anxiety associated using responses to the Patient Health Questionnaire–9 (PHQ-9) and Generalized Anxiety Disorder 7 (GAD-7). PHQ-9 has 9 questions related to depression status set [19,20], and GAD-7 is a 7-question instrument about general anxiety [21]. We evaluated the quality-adjusted life year (QALY) calculation in the EuroQol 5-Dimension 5-Level (EQ-5D-5L) instrument to assess the cost-effectiveness of the VCBT as a health care technology [22]. The EQ-5D-5L questions determine quality of life [22,23]. Health status is determined in five dimensions: degree of movement, personal management, normal activity, pain/discomfort, and anxiety/hiding.

Criteria Used to Define Therapeutic Response and Remission

To calculate responsiveness to VCBT treatment and remission rates after the VCBT, we used criteria employed by previous studies regarding the severity rating scales of the three disorders (Y-BOCS, PDSS, and LSAS). Regarding OCD, treatment response was defined as a 35% or greater reduction in the total Y-BOCS score, and remission was defined as a 12-month Y-BOCS≤14 [24]. Regarding PD, treatment response was defined as a 40% or greater reduction in total PDSS score, and remission was defined as a 12-month PDSS≤ 7 [25]. For SAD, treatment response was defined as a 31% or greater reduction in total LSAS score, and remission was defined as a 12-month LSAS≤35 [17].

Sources for Cost-Effectiveness

We calculated the total VCBT cost using the sum of the costs of implementing the intervention: (1) health care costs (¥3500-¥4800 × 16 sessions) and (2) costs of videoconferencing (license fee ¥1490 per month × 4 months in Webex (Cisco), ¥300 × 16 sessions in curon (MICIN, Inc) [26,27]. Note that in Japan, the cost of CBT in health care settings differs depending on whether it is performed by a doctor or jointly performed by a doctor and a nurse [28]. The cost-effectiveness threshold of the VCBT intervention was based on the willingness-to-pay (WTP) figure determined in a previous study (¥5 million) [29].

We did not assume that hardware would have to be newly purchased in order to access VCBT. This was because, as reported by the Ministry of Internal Affairs and Communications in 2017, ownership of information communication equipment in Japan was at 94.8% for mobile devices in general and 72.5% for PCs and because the penetration rate of information and communication equipment and the internet was at more than 80.9% for all households [30].

Statistical Analyses

Statistical analysis and reporting were performed in accord with the CONSORT-EHEALTH guidelines [31]. All statistical analyses were described in the statistical analysis plan, which was fixed before the database lock. All efficacy analyses were primarily based on the entire analytical dataset. Summary statistics were generated on all baseline variables with frequencies and proportions calculated on categorical data and means and standard deviations calculated on continuous variables.

The main analysis compared the baseline assessment scores with those obtained at the 12-month posttreatment follow-up. The differences were estimated using mixed-model analysis of variance (ANOVA) on all patients displaying symptoms in each scale (Y-BOCS, PDSS, LSAS, PHQ-9, and GAD-7), taking into account missing values, individual variance, and multiple measurement points.

Analysis of secondary outcomes was performed in an identical fashion to that of the primary analysis. To analyze cost-effectiveness using the EQ-5D-5L, QALY scores were estimated via area-under-the-curve analysis, which involved summing the areas of the distribution shapes for utility scores over the study period [22]. We calculated QALY summary statistics using the EQ-5D-5L data during the follow-up period complemented by multivariate imputation by chained equations (MICE) and last observation carried forward (LOCF). MICE was used as a guide for 100 completions [32].

The amount of change in QALY was calculated from the difference between QALY and the actually observed utility value assuming no change from the utility value of EQ-5D-5L at baseline. We calculated a summary statistic for the change in QALY and performed a paired *t* test. The method for calculating the change in QALY was as follows: QALY change amount = (baseline and end of treatment, 3 months, 6 months, 8 months, or area under the curve connecting the utility values including 12 months) – (effective value at each time point is baseline utility value and area under the curve assuming no change).

Cost-effectiveness of the VCBT was analyzed as follows. The additional consumption of health care resources was divided by the benefits (such as QALY) gained from the health care intervention to calculate an incremental cost-effectiveness ratio (ICER). When the ICER, such as cost per QALY, was less than a predetermined threshold, the intervention was considered cost-effective [33]. These thresholds were: (1) £20,000-£30,000 per QALY at the UK National Institute for Health and Care Excellence (NICE) [34], (2) $62,000 in the United States, and (3) ¥5 million in Japan [29]. The formulae used to calculate cost-effectiveness of VCBT, cost of VCBT, and WTP were as follows:

- Cost-effectiveness of VCBT = WTP − cost of VCBT
- Cost of VCBT = (videoconference system costs) + traditional CBT costs
- WTP = increased QALYs × threshold in Japan (¥5 million)

Calculated cost-effectiveness greater than one indicated that VCBT was a cost-effective intervention. WTP was calculated by multiplying the increase in QALY between baseline and 12-month follow-up after VCBT using the Japanese cost-effectiveness threshold (¥5 million). Incremental cost-effectiveness ratio per QALY was calculated by dividing the increase in QALY between baseline and 12-month follow-up after VCBT using total cost of VCBT.

Results

Participants

The sample comprised 4 males and 21 females, aged 20 to 54 years (mean 35.1 [SD 8.6] years) with 12 to 18 years of education (mean 14.72 [SD 1.90] years). Except for their principal diagnoses, participants' demographic and diagnostic data are described in Table 1, and the sampling procedure is illustrated in Figure 1.

Table 1. Participant clinical and demographic characteristics.

Characteristics	Overall (n=25)	OCD[a] (n=10)	PD[b] (n=7)	SAD[c] (n=8)
Age in years, mean (SD)	35.1 (8.6)	37.7 (6.9)	36.1 (9.3)	30.9 (9.4)
Gender (female), n (%)	21 (84)	8 (80)	7 (100)	6 (75)
Employed, n (%)	14 (56)	3 (12)	5 (71)	6 (75)
Pharmacotherapy (yes), n (%)	9 (36)	5 (50)	3 (43)	1 (13)
Comorbidity, n (%)				
Depression	3 (12)	1 (10)	0 (0)	2 (24)
Panic/agoraphobia	2 (11)	2 (20)	0 (0)	0 (0)
PTSD[d]	1 (4)	1 (10)	0 (0)	0 (0)
Alcohol dependence	1 (4)	0 (0)	0 (0)	1 (13)

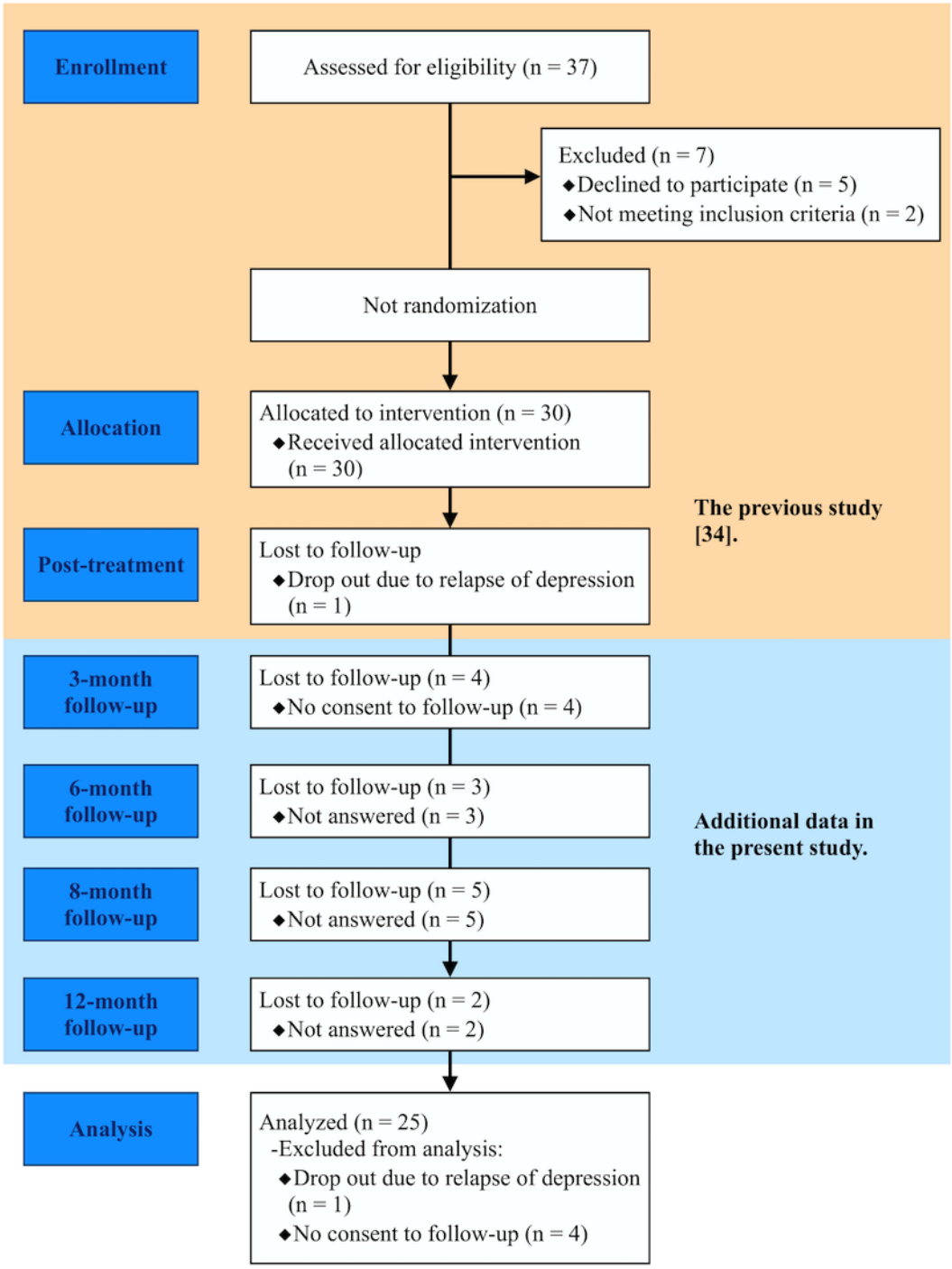
[a]OCD: obsessive-compulsive disorder.

[b]PD: panic disorder.

[c]SAD: social anxiety disorder.

[d]PTSD: posttraumatic stress disorder.

Figure 1. Participant flow.



Long-Term Effectiveness

Mixed-model ANOVA results regarding the long-term effectiveness of VCBT showed statistically significant improvement in participant symptoms (Table 2 and Figure 2). Changes in total mean scores between baseline and 12-month assessment were –4.1 on the Y-BOCS (F_1=4.45, P=.04), –4.4 on the PDSS (F_1=6.83, P=.01), and –30.9 on the LSAS (F_1=6.73, P=.01). Changes in the total PHQ-9 (depression) and GAD-7 (general anxiety) scores between baseline and 12-month follow-up assessment were –2.7 on the PHQ-9 (F_1=7.72, P=.007) and –3.0 on the GAD-7 (F_1= 7.09, P=.009).

Table 2. Mixed-model analysis of variance results on changes in participant symptomology.

Characteristics	n	Score mean (SD)	Min-max[a]	*P* value
Y-BOCS[b]	—	—	—	.04
Baseline	10	23.3 (6.5)	15 (36)	—
Posttreatment	10	17.1 (9.9)	2 (34)	—
3-month	10	19.4 (7.5)	9 (32)	—
6-month	10	18.6 (8.1)	7 (32)	—
12-month	10	19.2 (8.4)	8 (29)	—
PDSS[c]	—	—	—	.01
Baseline	7	8.9 (3.8)	5 (16)	—
Posttreatment	7	5.3 (6.7)	0 (19)	—
3-month	7	5.4 (4.9)	2 (13)	—
6-month	6	4.5(6.1)	0 (16)	—
12-month	6	4.5 (3.6)	0 (10)	—
LSAS[d]	—	—	—	.01
Baseline	8	96.6 (27.3)	53 (132)	—
Posttreatment	8	57.4 (34.7)	21 (128)	—
3-month	8	62.6 (34.4)	20 (112)	—
6-month	6	57.3 (31.9)	7 (85)	—
12-month	7	65.7 (43.8)	10 (118)	—
PHQ-9[e]	—	—	—	.007
Baseline	25	8.8 (6.2)	0 (23)	—
Posttreatment	25	6.8 (7.0)	0 (22)	—
3-month	25	7.2 (5.8)	0 (24)	—
6-month	22	6.6 (6.1)	0 (19)	—
12-month	23	6.1 (5.7)	0 (20)	—
GAD-7[f]	—	—	—	.009
Baseline	25	8.8 (5.3)	0 (20)	—
Posttreatment	25	5.5 (5.1)	0 (16)	—
3-month	25	7.2 (4.6)	0 (19)	—
6-month	25	6.3 (5.0)	0 (21)	—
12-month	23	5.8 (4.5)	0 (14)	—

[a]min-max: minimum to maximum.

[b]Y-BOCS: Yale-Brown Obsessive-Compulsive Scale.

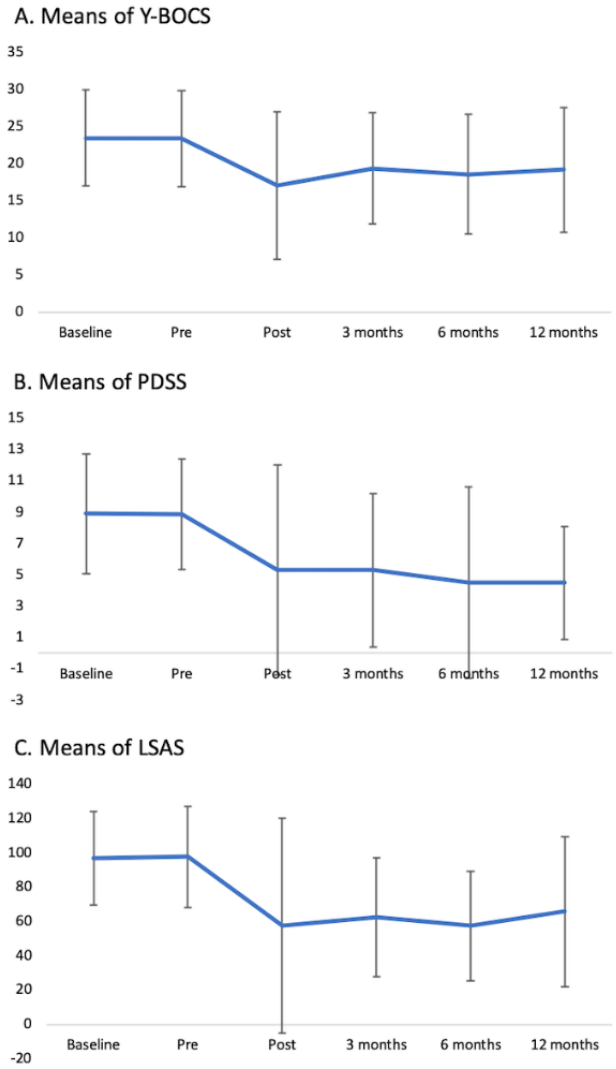
[c]PDSS: Panic Disorder Severity Scale.

[d]LSAS: Livobitz Social Anxiety Scale.

[e]PHQ-9: Patient Health Questionnaire–9.

[f]GAD-7: Generalized Anxiety Disorder–7.

Figure 2. Participants' changes in symptomology.



To investigate the predictive effects that symptoms of patients' depression at pretreatment may have had on the treatment response change posttreatment, multiple regression analyses in simultaneous forced entry were performed. The treatment response percentage change was set as a dependent variable in multiple regression analyses. We set depressive symptoms due to PHQ-9 as independent variables. The treatment response percentage change was calculated by dividing the total baseline score with the score difference between baseline and 12-month. The treatment response percentage change in this study was the decline in baseline Y-BOCS, PDSS, or LSAS score.

The degree of change (in percentages) in the treatment response was analyzed as a continuous variable. Statistical analysis was performed using SPSS Statistics version 24.00 (IBM Corp). Multiple regression analysis showed that the effects of depression on therapeutic response rates were not significant across the data (β=–1.74, adjusted R^2=.13, SE 25.29, P=.053, VIF (variance inflation factor)=1.00), OCD (β=–1.60, adjusted R^2=.24, SE 19.46, P=.16), PD (β=–0.41, adjusted R^2=.25, SE 38.82, P=.96), and SAD (β=–0.73, adjusted R^2=.18, SE 24.02, P=.77).

Therapeutic Response and Remission Rates

At the 12-month follow-up assessment, treatment response rate was 32% (8/25) and remission rate was 40% (10/25; Table 3).

Table 3. Participant response and remission rates at each follow-up end point.

Characteristics	Overall (n=25), n (%)	OCD[a] (n=10), n (%)	PD[b] (n=7), n (%)	SAD[c] (n=8), n (%)
Response				
Posttreatment	12 (48)	4 (40)	4 (57)	4 (50)
3-month	10 (40)	2 (20)	5 (71)	3 (38)
6-month	8 (32)	2 (20)	3 (43)	3 (38)
12-month	8 (32)	2 (20)	2 (29)	4 (50)
Remission				
Posttreatment	12 (48)	4 (40)	6 (86)	2 (25)
3-month	11 (44)	4 (40)	5 (71)	2 (25)
6-month	10 (40)	3 (30)	5 (71)	2 (25)
12-month	10 (40)	3 (30)	5 (71)	2 (25)

[a]OCD: obsessive-compulsive disorder.

[b]PD: panic disorder.

[c]SAD: social anxiety disorder.

Cost-Effectiveness

Table 4 shows the EQ-5D-5L index for each end point. The 1-year converted QALY score from baseline to 12 months posttreatment was 0.7469 (SE 0.0353, 95% CI 0.6728-0.821), and the change between baseline and 12-month follow-up assessment was 0.0379 (SE 0.01; Table 5). Figure 3 shows the QALY calculated from the EQ-5D-5L between baseline and 12-month follow-up assessment. There was a significant increase of 0.038 (95% CI 0.0085-0.0674, *P*=.02) in complete cases. The results on the data supplemented with missing values are shown in Table 6. The WTP threshold was ¥189,500 because the 0.0379 score in the QALYs increased after the intervention. The health care costs including the VCBT accounted for the CBT health care costs (¥56,000-¥76,000), and annual licensing fees per patient for the videoconferencing system (¥4800-¥5960) was ¥60,800 to ¥81,960 (Table 7). Thus, we concluded that the VCBT was a cost-effective intervention because VCBT costs were below the threshold set for the cost-effectiveness analysis.

Table 4. EuroQol 5-Dimension 5-Level index each end point.

Characteristics	n	Mean	SD	SE
Complete case				
Baseline	25	0.7206	0.14	—
Posttreatment	25	0.7677	0.20	—
3-month	25	0.7350	0.17	—
6-month	22	0.7207	0.24	—
8-month	20	0.7760	0.15	—
12-month	23	0.7503	0.15	—
LOCF[a]				
6-month	25	0.7342	0.23	—
8-month	25	0.7669	0.15	—
12-month	25	0.7530	0.15	—
MICE[b]				
6-month	25	0.7075	—	0.05
8-month	25	0.7651	—	0.03
12-month	25	0.7564	—	0.03

[a]LOCF: last observation carried forward.

[b]MICE: multivariate imputation by chained equations.

Table 5. Paired *t* test results on change of quality-adjusted life years at 12 months after baseline.

Characteristics	n	Mean	SE	95% CI	*P* value
Complete cases	19	0.0379	0.01	0.0085-0.0674	.02
LOCF[a]	25	0.0214	0.01	0.0067-0.0495	.13
MICE[b]	25	0.0187	0.01	0.0093-0.0466	.19

[a]LOCF: last observation carried forward.

[b]MICE: multivariate imputation by chained equations.

Figure 3. The quality-adjusted life years (QALYs) observed at follow-up and QALY in complete cases. Note: Estimated QALYs was calculated in terms of effective value without videoconference-delivered cognitive behavioral therapy at each time point as a baseline, with the area under the curve assuming no change. Increased QALY was calculated as the difference between the measured utility value and the estimated QALY.

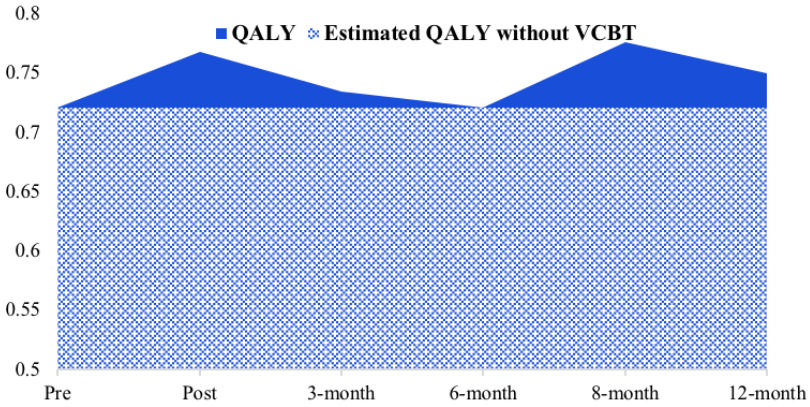


Table 6. Quality-adjusted life years at 12 months after videoconference-delivered cognitive behavioral therapy.

Characteristics	n	Mean	SE	95% CI
Complete cases	19	0.7469	0.04	0.6728-0.8210
LOCF[a]	25	0.7420	0.03	0.6839-0.8001
MICE[b]	25	0.8343	0.04	0.7565-0.9121

[a]LOCF: last observation carried forward.

[b]MICE: multivariate imputation by chained equations.

Table 7. Results of a cost-utility analysis.

Characteristics	Value
Cost for each service per patient (¥)	
CBT[a] by a nurse	¥56,000
CBT by a medical doctor	¥76,000
Videoconferencing in Webex	¥5960
Videoconferencing in Curon	¥4800
Total cost	¥60,800-¥81,960
QALY[b]	
Complete case	0.7469
LOCF[c]	0.742
MICE[d]	0.8343
Incremental benefit, QALY gain	
Complete case	0.0379
LOCF	0.0214
MICE	0.0187
Incremental cost-effectiveness ratio per QALY	
Complete case	¥1,604,222 to ¥2,162,533
LOCF	¥2,841,122 to ¥3,829,907
MICE	¥3,251,337 to ¥4,382,888
Willingness to pay = ¥5 million per QALY	
Complete case	¥189,500
LOCF	¥107,000
MICE	¥93,500

[a]CBT: cognitive behavioral therapy.

[b]QALY: quality adjusted life year.

[c]LOCF: last observation carried forward.

[d]MICE: multivariate imputation by chained equations.

Discussion

Principal Findings

We investigated the long-term effectiveness and cost-effectiveness of VCBT in 25 patients with OCD, PD, or SAD in a 12-month observational study. The principal symptomology of OCD, PD, and SAD significantly decreased and the QALY significantly improved. The therapeutic response rate was 32% (8/25) and remission rate was 40% (10/25) at the 12-month postintervention follow-up assessment. The total cost of providing VCBT was ¥60,800 to ¥81,960 per patient; in contrast, the threshold using WTP was ¥189,500. Therefore, our results suggested that VCBT was a cost-effective intervention for this sample of patients with OCD, PD, or SAD in Japan.

Long-Term Effectiveness of Videoconference-Delivered Cognitive Behavioral Therapy

In a previous study on VCBT provided to 10 adult OCD patients, 2 patient scores were below the Y-BOCS cutoff (<14) after treatment, but just one patient was below the cutoff 3 months later [10]. There was a trend toward increased OCD symptoms at 3 months' postintervention [10]. A randomized controlled trial (RCT) of VCBT provided to OCD patients aged 7 to 16 years indicated that continued improvement was observed in the symptoms until 3 months after treatment [35]. In that study [35], one of the two patients who evidenced remission before and after the treatment was still in remission 6 months' postintervention, whereas the other patient presented worse symptoms. This study provides observational results from the end point of the VCBT for 12 months, which extends the findings of previous research. In other words, as the amount of time after the intervention increased, OCD symptoms apparently increased and the proportion of remissions apparently decreased from 40% (4/10) immediately after VCBT to 30% (3/10) at 3 months later and 20% (2/10) at the 6-month and 12-month follow-up assessments. The results of this study are consistent with previous studies that the remission rate decreases with time [7,30]. When patient symptoms increase, they might access a self-help program or attend regular support sessions to help

prevent symptom relapse [36]. In a survey of adolescents with OCD, satisfaction with support sessions was universal [37].

In a study that provided VCBT to 11 adult PD patients, 82% (9/11) had improved symptoms after the intervention and 91% (10/11) had improved symptoms after 6 months and no panic attacks [38]. This study extended the examination of the long-term efficacy of VCBT in patients with PD to 12 months and found that it was effective for 85% (6/7) after treatment, and it held steady at 71% (5/7) after 3, 6, and 12 months. However, although VCBT has demonstrated its long-term efficacy, there is some indication that panic symptoms might relapse over time [25].

In a study of VCBT in 24 adult patients with SAD, 54% (13/24) experienced remission after treatment, and symptoms that had decreased were maintained at that lower level 6 months later [39]. Our results were similar to that study in that the patients who achieved remission after VCBT seemed to continue in remission for 6 or 12 months (both 2/8, 25%). The same two patients exhibited remission at any point during the 12 months. We gradually lost contact during the observation period with 2 of the 4 patients who had exhibited remission after the VCBT. Therefore, when interpreting the long-term symptom-improving effects, it might be important to consider the course of remission rather than the overall remission rate during the study period.

Cost-Effectiveness of Videoconference-Delivered Cognitive Behavioral Therapy

Several studies have reported that internet-delivered cognitive behavioral therapy (ICBT) provided to patients with depression saved on direct medical costs more than providing just the usual care [40,41]. In an RCT conducted in Spain [41], providing ICBT to patients with depression was more cost-effective than 12 months of treatment restricted to usual care: €6381 for therapist-guided ICBT and €11,390 for nonguided ICBT. On the other hand, an RCT of ICBT aimed at preventing recurrent depression found that the average cost after 24 months was not significantly different between the ICBT group ($8298) and the usual care group ($7296) [42]. A study of face-to-face CBT in 469 participants with depression suggested that the incremental cost-effectiveness ratio was £5374 per QALY gain [43], below the threshold £20,000 to £30,000 at NICE [34]. That study's result was consistent with our result: ¥1,604,222 to ¥2,162,533 per QALY gain (£11,459 to £15,447; calculated as ¥100=£140), below the threshold of ¥5 million in Japan [29]. Hence, CBT for depression and anxiety disorders was cost-effective whether it was face-to-face or internet intervention, with or without videoconferencing.

This study provides the world's first empirical knowledge about the cost-effectiveness of VCBT. VCBT costs totaled ¥60,800 to ¥81,960, which was far below the ¥189,500 threshold based on WTP calculated using the QALY. In other words, under the Japanese insurance system in 2018 [28], VCBT was a cost-effective treatment approach. We determined that ¥100 was approximately $110, €120, and £140. The VCBT costs were then determined to be $553 to $745, €507 to €683, and £434 to £585 and the threshold of WTP was $1723, €1579, £1354.

Limitations and Future Research

This study has some limitations. First, there was no statistical control over the relationship between VCBT and pharmacological therapy during our previous trial and this follow-up study. Studies have suggested that combining therapeutic approaches with drug therapy is particularly effective in panic disorder prognoses [44]. Future studies should include a controlled design that accounts for drug therapy and combination therapy. Second, we did not account for the effects of support provided to the participants during the observation period after the VCBT. Patients who continue to use antidepressants after remission were known to have a lower recurrence rate than those who discontinued prematurely [45]. Third, there was no usual care group to contrast with the VCBT group as a control in the cost-effectiveness analysis. We examined the cost-effectiveness of VCBT based on a white paper on the health care costs of patients with anxiety disorders in Japan [28], and, therefore, future research should employ actual observations and data. Fourth, during some observation periods (eg, 6 months or 12 months posttreatment), we lacked data on participants who exhibited significant symptom improvements immediately after treatment. Therefore, the results should be interpreted with caution. Fifth, a small sample size was used in this study, and there was no comparison group. Future studies should use a large sample and employ RCTs. Sixth, participants recruited in this study tended to be living in their own local areas far from our hospital without having face-to-face CBT, and they had relatively long duration of untreated illness before CBT. In future, VCBT cost-effectiveness studies for patients at the early onset stage of the disorders in primary care settings will be required.

Conclusion

Our results suggest that VCBT for patients with OCD, PD, and SAD was effective in improving symptoms over 12 months and was a cost-effective approach in Japan.

Acknowledgments

We express our gratitude and respect for the dedication and contribution of research participants. This study was supported by the Japan Society for the Promotion of Science KAKENHI Grant-in-Aid for Scientific Research C (Nos. 18K03130 and 18K17313). The funding sources had no role in the design and conduct of the study.

Authors' Contributions

KM contributed to design of research, data collection and analysis, and the development of the manuscript. SH contributed to data collection. KN contributed to the design and statistical analysis in this study. CS contributed to management of this study.

AN contributed to data collection. ES contributed to the drafting and planning of this study and provided interpretations based on the results.

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington: American Psychiatric Association; 2013.
2. Eisen JL, Sibrava NJ, Boisseau CL, Mancebo MC, Stout RL, Pinto A, et al. Five-year course of obsessive-compulsive disorder: predictors of remission and relapse. J Clin Psychiatry 2013 Mar;74(3):233-239 [FREE Full text] [doi: 10.4088/JCP.12m07657] [Medline: 23561228]
3. Spinhoven P, Batelaan N, Rhebergen D, van Balkom A, Schoevers R, Penninx BW. Prediction of 6-yr symptom course trajectories of anxiety disorders by diagnostic, clinical and psychological variables. J Anxiety Disord 2016 Dec;44:92-101. [doi: 10.1016/j.janxdis.2016.10.011] [Medline: 27842240]
4. Scholten WD, Batelaan NM, Penninx BW, van Balkom A, Smit JH, Schoevers R, et al. Diagnostic instability of recurrence and the impact on recurrence rates in depressive and anxiety disorders. J Affect Disord 2016 May;195:185-190. [doi: 10.1016/j.jad.2016.02.025] [Medline: 26896812]
5. Ministry of Health, Labour and Welfare. Health Insurance URL: https://www.mhlw.go.jp/bunya/iryouhoken/iryouhoken01/dl/01_eng.pdf [accessed 2020-04-09]
6. Matsumoto K, Sutoh C, Asano K, Seki Y, Urao Y, Yokoo M, et al. Internet-based cognitive behavioral therapy with real-time therapist support via videoconference for patients with obsessive-compulsive disorder, panic disorder, and social anxiety disorder: pilot single-arm trial. J Med Internet Res 2018 Dec 17;20(12):e12091 [FREE Full text] [doi: 10.2196/12091] [Medline: 30559094]
7. Moreno FA, Chong J, Dumbauld J, Humke M, Byreddy S. Use of standard Webcam and Internet equipment for telepsychiatry treatment of depression among underserved Hispanics. Psychiatr Serv 2012 Dec;63(12):1213-1217. [doi: 10.1176/appi.ps.201100274] [Medline: 23026854]
8. De Las Cuevas C, Arredondo MT, Cabrera MF, Sulzenbacher H, Meise U. Randomized clinical trial of telepsychiatry through videoconference versus face-to-face conventional psychiatric treatment. Telemed J E Health 2006 Jun;12(3):341-350. [doi: 10.1089/tmj.2006.12.341] [Medline: 16796502]
9. Stubbings DR, Rees CS, Roberts LD, Kane RT. Comparing in-person to videoconference-based cognitive behavioral therapy for mood and anxiety disorders: randomized controlled trial. J Med Internet Res 2013;15(11):e258 [FREE Full text] [doi: 10.2196/jmir.2564] [Medline: 24252663]
10. Vogel PA, Solem S, Hagen K, Moen EM, Launes G, Håland AT, et al. A pilot randomized controlled trial of videoconference-assisted treatment for obsessive-compulsive disorder. Behav Res Ther 2014 Oct 30;63C:162-168. [doi: 10.1016/j.brat.2014.10.007] [Medline: 25461792]
11. 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 [FREE Full text] [Medline: 9881538]
12. Muramatsu K, Miyaoka H, Kamijima K, Muramatsu Y, Yoshida M, Otsubo T, et al. The patient health questionnaire, Japanese version: validity according to the mini-international neuropsychiatric interview-plus. Psychol Rep 2007 Dec;101(3 Pt 1):952-960. [doi: 10.2466/pr0.101.3.952-960] [Medline: 18232454]
13. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, et al. The Yale-Brown Obsessive Compulsive Scale. I. development, use, and reliability. Arch Gen Psychiatry 1989 Nov;46(11):1006-1011. [doi: 10.1001/archpsyc.1989.01810110048007] [Medline: 2684084]
14. Hamagaki S, Takagi S, Urushihara Y, Ishisaka Y, Matsumoto M. [Development and use of the Japanese version of the self-report Yale-Brown Obsessive Compulsive Scale]. Seishin Shinkeigaku Zasshi 1999;101(2):152-168. [Medline: 10375975]
15. Houck PR, Spiegel DA, Shear MK, Rucci P. Reliability of the self-report version of the panic disorder severity scale. Depress Anxiety 2002;15(4):183-185. [doi: 10.1002/da.10049] [Medline: 12112724]
16. Katagami M. The self-report version of the Panic Disorder Severity Scale: reliability and validity of the Japanese version. Jpn J Psychosom Med 2007;47:331-338. [doi: 10.15064/jjpm.47.5_331]
17. Liebowitz M. Social phobia. Mod Probl Pharmacopsychiatry 1987;22:141-173. [doi: 10.1159/000414022] [Medline: 2885745]
18. Asakura S, Inoue S, Sasaki F, Sasaki Y, Kitagawa N, Inoue T. Consideration of validity and reliability for Liebowitz Social Anxiety Scale (LSAS) Japanese version. Clin Psychiatry 2002;177:1084.
19. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]

20. Muramatsu K. Patient Health Questionnaire (PHQ-9, PHQ-15) Japanese version and Generalized Anxiety Disorder-7 Japanese version up to date. Stud Clin Psychol 2014;7:35-39.
21. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
22. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
23. Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ 2002 Jun;11(4):341-353. [doi: 10.1002/hec.673] [Medline: 12007165]
24. Farris SG, McLean CP, Van Meter PE, Simpson HB, Foa EB. Treatment response, symptom remission, and wellness in obsessive-compulsive disorder. J Clin Psychiatry 2013 Jul;74(7):685-690 [FREE Full text] [doi: 10.4088/JCP.12m07789] [Medline: 23945445]
25. Furukawa TA, Katherine SM, Barlow DH, Gorman JM, Woods SW, Money R, et al. Evidence-based guidelines for interpretation of the Panic Disorder Severity Scale. Depress Anxiety 2009;26(10):922-929 [FREE Full text] [doi: 10.1002/da.20532] [Medline: 19006198]
26. CISCO. Webex URL: https://www.webex.com/ [accessed 2019-11-20]
27. Micin Inc. Curon URL: https://curon.co/ [accessed 2019-11-20]
28. 61st Social Security Council Medical Committee in Ministry of Health, Labour and Welfare. Summary of 2018 Medical Fee Revision URL: https://www.mhlw.go.jp/file/05-Shingikai-12601000-Seisakutoukatsukan-Sanjikanshitsu_Shakaihoshoutantou/0000203227.pdf [accessed 2019-10-11]
29. Shiroiwa T, Sung Y, Fukuda T, Lang H, Bae S, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ 2010 Apr;19(4):422-437. [doi: 10.1002/hec.1481] [Medline: 19382128]
30. Ministry of Internal Affairs and Communications. Information and Communications in Japan: white paper 2019 URL: https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2019/2019-index.html [accessed 2019-12-18]
31. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
32. Wang B, Fang Y, Jin M. Combining type-III analyses from multiple imputations. 2014. URL: https://support.sas.com/resources/papers/proceedings14/1543-2014.pdf [accessed 2020-04-09]
33. Towse A. What is NICE's threshold? An external view. In: Devlin N, Tows A, editors. Cost Effectiveness Thresholds: Economic and Ethical Issues. Office for Health Economica: King's Fund; 2002.
34. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ 2004 May;13(5):437-452. [doi: 10.1002/hec.864] [Medline: 15127424]
35. Storch E, Caporino N, Morgan J, Lewin A, Rojas A, Brauer L, et al. Preliminary investigation of web-camera delivered cognitive-behavioral therapy for youth with obsessive-compulsive disorder. Psychiatry Res 2011 Oct 30;189(3):407-412. [doi: 10.1016/j.psychres.2011.05.047] [Medline: 21684018]
36. Andersson E, Steneby S, Karlsson K, Ljótsson B, Hedman E, Enander J, et al. Long-term efficacy of Internet-based cognitive behavior therapy for obsessive-compulsive disorder with or without booster: a randomized controlled trial. Psychol Med 2014 Oct;44(13):2877-2887. [doi: 10.1017/S0033291714000543] [Medline: 25066102]
37. Negreiros J, Selles RR, Lin S, Belschner L, Stewart SE. Cognitive-behavioral therapy booster treatment in pediatric obsessive-compulsive disorder: a utilization assessment pilot study. Ann Clin Psychiatry 2019 Aug;31(3):179-191. [Medline: 31369657]
38. Bouchard S, Paquin B, Payeur R, Allard M, Rivard V, Fournier T, et al. Delivering cognitive-behavior therapy for panic disorder with agoraphobia in videoconference. Telemed J E Health 2004;10(1):13-25. [doi: 10.1089/153056204773644535] [Medline: 15104911]
39. Yuen E, Herbert J, Forman E, Goetter EM, Juarascio A, Rabin S, et al. Acceptance based behavior therapy for social anxiety disorder through videoconferencing. J Anxiety Disord 2013 May;27(4):389-397. [doi: 10.1016/j.janxdis.2013.03.002] [Medline: 23764124]
40. 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]
41. Romero-Sanchiz P, Nogueira-Arjona R, García-Ruiz A, Luciano J, García Campayo J, Gili M, et al. Economic evaluation of a guided and unguided internet-based CBT intervention for major depression: Results from a multi-center, three-armed randomized controlled trial conducted in primary care. PLoS One 2017;12(2):e0172741 [FREE Full text] [doi: 10.1371/journal.pone.0172741] [Medline: 28241025]
42. 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]
43. Wiles NJ, Thomas L, Turner N, Garfield K, Kounali D, Campbell J, et al. Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care:

follow-up of the CoBalT randomised controlled trial. Lancet Psychiatry 2016 Feb;3(2):137-144 [FREE Full text] [doi: 10.1016/S2215-0366(15)00495-2] [Medline: 26777773]
44. Furukawa TA, Watanabe N, Churchill R. Psychotherapy plus antidepressant for panic disorder with or without agoraphobia: systematic review. Br J Psychiatry 2006 Apr;188:305-312 [FREE Full text] [doi: 10.1192/bjp.188.4.305] [Medline: 16582055]
45. Betelann N, Bosman R. Risk of relapse after antidepressant discontinuation in anxiety disorders, obsessive-compulsive disorder, and post-traumatic stress disorder: systematic review and meta-analysis of relapse prevention trials. BMJ 2017 Sep 25;358:j4461 [FREE Full text] [doi: 10.1136/bmj.j4461] [Medline: 28947609]

Abbreviations

ANOVA: analysis of variance
CBT: cognitive behavioral therapy
EQ-5D-5L: EuroQol 5-Dimension 5-Level
GAD-7: Generalized Anxiety Disorder–7
ICBT: internet-delivered cognitive behavioral therapy
ICER: incremental cost-effectiveness ratio
LOCF: last observation carried forward
LSAS: Liebowitz Social Anxiety Scale
MICE: multivariate imputation by chained equations
NICE: National Institute for Health and Care Excellence
OCD: obsessive-complusive disorder
PD: panic disorder
PDSS: Panic Disorder Severity Scale
PHQ-9: Patient Health Questionnaire–9
QALY: quality-adjusted life year
RCT: randomized controlled trial
SAD: social anxiety disorder
VCBT: videoconference-delivered cognitive behavioral therapy
WTP: willingness-to-pay
Y-BOCS: Yale-Brown Obsessive-Compulsive Scale

Edited by G Eysenbach; submitted 22.11.19; peer-reviewed by K Mathiasen, R Nogueira-Arjona, T Muto; comments to author 16.12.19; revised version received 10.01.20; accepted 25.03.20; published 23.04.20.

Please cite as:
Matsumoto K, Hamatani S, Nagai K, Sutoh C, Nakagawa A, Shimizu E
Long-Term Effectiveness and Cost-Effectiveness of Videoconference-Delivered Cognitive Behavioral Therapy for Obsessive-Compulsive Disorder, Panic Disorder, and Social Anxiety Disorder in Japan: One-Year Follow-Up of a Single-Arm Trial
JMIR Ment Health 2020;7(4):e17157
URL: http://mental.jmir.org/2020/4/e17157/
doi:10.2196/17157
PMID:32324150

Original Paper

How Contextual Constraints Shape Midcareer High School Teachers' Stress Management and Use of Digital Support Tools: Qualitative Study

Julia B Manning[1], BSc, MCOptom; Ann Blandford[2], CEng, PhD; Julian Edbrooke-Childs[3], PhD; Paul Marshall[4], DPhil

[1]University College London Interaction Centre, London, United Kingdom

[2]Institute of Healthcare Engineering, University College London, London, United Kingdom

[3]Evidence-based Practice Unit, Anna Freud Centre and University College London, London, United Kingdom

[4]Department of Computer Science, University of Bristol, London, United Kingdom

Corresponding Author:
Julia B Manning, BSc, MCOptom
University College London Interaction Centre
66-72 Gower Street
London, WC1E 6EA
United Kingdom
Phone: 44 7973312358
Email: j.manning.17@ucl.ac.uk

Abstract

Background: Persistent psychosocial stress is endemic in the modern workplace, including among midcareer high school (secondary comprehensive) teachers in England. Understanding contextual influences on teachers' self-management of stress along with their use of digital health technologies could provide important insights into creating more usable and accessible stress support interventions.

Objective: The aim of this study was to investigate the constraints on stress management and prevention among teachers in the school environment and how this shapes the use of digitally enabled stress management tools.

Methods: Semistructured interviews were conducted with 14 teachers from southern England. The interviews were analyzed using thematic analysis.

Results: Teachers were unanimous in their recognition of workplace stress, describing physical (such as isolation and scheduling) and cultural (such as stigma and individualism) aspects in the workplace context, which influence their ability to manage stress. A total of 12 participants engaged with technology to self-manage their physical or psychological well-being, with more than half of the participants using consumer wearables, but Web-based or smartphone apps were rarely accessed in school. However, digital well-being interventions recommended by school leaders could potentially be trusted and adopted.

Conclusions: The findings from this study bring together both the important cultural and physical contextual constraints on the ability of midcareer high school teachers to manage workplace stress. This study highlights correlates of stress and offers initial insight into how digital health interventions are currently being used to help with stress, both within and outside high schools. The findings add another step toward designing tailored digital stress support for teachers.

(JMIR Ment Health 2020;7(4):e15416) doi:10.2196/15416

KEYWORDS

school teachers; stress; health; self-management; computers; technology; qualitative research; secondary schools; wearable devices; mobile applications; education

Introduction

Background

Teaching is one of the most stressful professions. Surveys have shown teaching to be in the top 6 (out of 26) professions that report the worst scores for physical health, psychological well-being, and job satisfaction [1]. Excessive, chronic stress is acknowledged as a condition of mental or emotional strain or tension, which can have psychological, behavioral, and physiological effects [2]. "Teacher stress" was first specifically described by Kyriacou [3] as "a response syndrome of negative effects (such as anger or depression) resulting from the general teacher's job". Teachers' experiences of stress at work have received extensive coverage in the literature [4-8], with stress sources recently summarized as the following: (1) teaching unmotivated students, (2) maintaining discipline, (3) time pressures and workload, (4) coping with change, (5) being evaluated by others, (6) dealings with colleagues, (7) self - esteem and status, (8) administration and management, (9) role conflict and ambiguity, and (10) poor working conditions [9]. Some of these sources reflect contextual factors that are intrinsic to the job and shape the method of teaching (eg, classroom setting, timetabling, and instructing), which have also been referred to as the *school climate* [10,11]. Stress coping strategies have been categorized as physiological (eg, relaxation, meditation, and aerobic activity), situational (eg, changing personal reaction or altering work environment), or cognitive (eg, controlling emotions, problem solving, or time management) [12]. Cognitive strategies have been shown to be used more than situational strategies [13,14], but physiological interventions, including mindfulness, have been growing in popularity [8,15,16].

Stress management has often been analyzed at the level of intervention. Stress interventions for teachers have been categorized by the level of mediation: (1) organizational, focused on the organization's culture; (2) organization-individual interface, including building workplace relations; or (3) individual, where people are taught practices to manage stress [17]. The following fourth intervention categorization has subsequently been added: classroom level, influencing student functioning and behavior [18]. Studies to date indicate that organizational health interventions have the most potential to reduce general teachers' work-related stress [19]. Yet, this seems to have only been verified in the US primary (elementary) school system, not in other countries (eg, England) and never in high schools. In addition, although some organizational contextual factors such as school leadership, workplace culture, and teacher autonomy have been identified as playing a critical role in managing stress, it is noted that they are in the purview of *schools and employing bodies* and not under the control of an individual's authority or management [20].

A systematic review of studies on the effectiveness of taught interventions aimed at reducing general teacher burnout (ie, stress-induced work absence) found that in the 11 out of 23 controlled studies which included high school–level educators, cognitive behavioral therapy (CBT), mindfulness, and social support were the interventions that showed an effect, albeit small [21]. The lack of tailoring of these interventions to teachers' work and context is suggested as the reason for the effect size being small, as tailored interventions showed more positive results. Tailoring included reflecting classroom situations and teacher-student relationships [22-24]. Many approaches to well-being, described as *whole school* do not include the staff's own well-being [25-27], but there is an established literature on the link between teachers' well-being and work-related outcomes [28-30], which some schools are beginning to recognize [31,32].

Technology has enabled stress reduction techniques to become digitized with digital health interventions (DHIs), generally known as electronic health (eHealth), and arguably more accessible, self-administered, and connected [33]. There is a growing body of studies detailing DHI use in the workplace. A meta-analysis of 21 randomized controlled trials (RCTs) on psychological interventions in the workplace utilizing Web-based programs (of which 3 used mobile apps) has shown improvements in employees' well-being and work effectiveness [34]. In another review of 23 eHealth RCTs, improvements have been seen in mental health and stress symptoms [35], although it was observed that similar to nondigital approaches, the best outcomes depend on providing the right type of intervention to the correct target population. Interestingly, introducing activity tracking or tools (eg, exercise videos, Nintendo Wii) intended to increase exercise levels among nurses did not result in lifestyle behavior change, but it did serve as a stress reliever [36]. Wearable devices have received little attention in the workplace so far. One study on activity with a small sample found that among office workers, wearable feedback did not affect sitting time and that there were issues with accuracy, battery life, and real-time feedback [37]. Another study with office workers found that on applying machine learning to the collected heart and activity data, it was possible to automate a degree of mood recognition, including stress, and therefore make some objective predictions, which posited that this could enable the employer to change the work environment accordingly [38]. However, the accuracy of wearables in stress detection in real-world situations remains as issue, as does the lack of agreement on a standard stress measurement method [39,40]. Only one tailored eHealth stress reduction study by Ebert et al in Germany [41] appears to have been conducted in education: an internet-based Problem Solving Therapy (PST, a form of CBT) adapted for teachers. Averaging 19 years of teaching experience, participants had elevated symptoms of depression. The PST resulted in a reduction in their perceived stress but it is unclear whether the study resulted in any systematic implementation. In the literature, there is little evidence of the adoption of DHIs for teachers' stress management in schools.

Teachers in the middle of their careers are at stages three and four in Huberman's widely accepted 5 stages of the teacher career development model. This includes those with 7 to 30 years of practice with roles including coordinators, heads of department, and in personal development [42]. They are more likely to be under pressure because of the demands both in and outside the job [43], but as *survivors* of the early years, they should also have management strategies in place [44]. A 2012

German study of nearly 2000 teachers found that midcareer teachers have more tasks and invest more effort than colleagues who are at the beginning or end of their careers [45]. Therefore, the combination of pressure, adaption, and effort seen in senior teachers means it is appropriate to focus the investigation on their stress management, particularly in schools where there is already an acknowledgment of the philosophies of workplace health [46]. Tailoring technology to real-world situations and teachers' existing behavior needs to be understood if we intend to align digital support in a meaningful way [34,35]. The literature on teacher stress is extensive but not conclusive.

Objectives

There are gaps in understanding how the educational context influences the strategies that midcareer high school teachers are actually able to use to manage stress. Similarly, we do not know to what extent teachers have already adopted technology as a stress intervention or prevention tool and what the opportunities are for expansion.

A more nuanced understanding of these teachers' real-world situation is required if we are to facilitate techniques for the reduction of stress, whether at the individual or organizational level. How does the school organizational context constrain individual teachers' stress management strategies? Therefore, the aims of this study were to engage with midcareer high school (secondary comprehensive) teachers to understand (1) how workplace context influences stress management and (2) how this affects the current use of digital health technology, to ascertain the potential for technology support for stress reduction.

Methods

Study Design

The methodological orientation of this study was experiential qualitative research [47]. Exploring and understanding participants' strategies in their everyday work context was desired so that these data might inform a future qualitative inquiry into the potential for integrating technological support for stress reduction among senior teachers.

Teacher Recruitment

The initial selection of participants was purposive so that the target population of high school teachers in midcareer roles was secured. The inclusion criteria were to be a high school teacher in a management position, eg, in the senior leadership team, head of department or head of house, or a teacher who had previously held a senior role. Willingness to talk about stress was specified in the information sheet. No prior individual training in stress interventions was stipulated, but their school had to have demonstrated an awareness of the importance of well-being, such as through having a school lead on well-being. Staff from 5 different schools in the midlands, London, and southern England, who met the inclusion criteria, were recruited using the researcher's networks (with the head teacher's permission sought for 2 on-site sets of interviews). None of the participants had previously met the researcher. All volunteers were sent a comprehensive information sheet and consent form to read when they had indicated that they were willing to be interviewed. The consent and information forms specified the purpose of the interview, the interviewer's department and university, and that the interviews would contribute to a study being undertaken as part of a PhD.

School-facilitated interviews were requested to take place during the school day to not cause teaching staff the inconvenience of having to stay on after hours. However, several staff members volunteered to be interviewed on a Friday afternoon after classes had finished. A printout of the information document was also given to participants at the start of the interview to ensure they had a physical copy to take away. This included information on where they could obtain free and confidential advice should they have felt in need of support after the interview had finished. This research was approved by the University College London (UCL) Research Ethics Committee approval ID number: UCLIC/1718/013/Staff Marshall/Manning.

Interview Script Development

The interview guide was informed by the literature on stress in the school workplace as well as previous Education Support Partnership stress questionnaires [48] and Health and Safety Executive research [49,50]. The interview took a semistructured qualitative approach with questions devised to understand (1) how workplace context influences stress management and (2) how this affects the current use of digital health technology, to ascertain the potential for technology support for stress reduction.

The interview had been pilot tested with 2 teachers independently, an assistant head and a head of department from 2 separate schools. The data from these interviews were only used to refine the interview schedule and not used in the thematic analysis. The whole interview schedule is presented in Multimedia Appendix 1. For the purposes of this paper, the focus was on the influence of the workplace context and insights into the use of digital health tools. None of the interviews were terminated prematurely or repeated. They were timed to last for 30 min, with the shortest interview lasting 22 min and the longest interview lasting 43 min.

Data Analysis

All interviews were transcribed verbatim and analyzed using inductive thematic analysis. NVivo version 12 (QSR International Ltd) was used for the organization and development of codes and categories, and SimpleMind (version 1.22) was used for the mapping of relationships and themes by the researcher (JM). Transcripts were not returned to participants for comments or corrections (1) as the researcher did not want to add a time burden on the participant, (2) as the evidence is that it adds little to the accuracy of the transcript [51], and (3) to retain the responses as precisely as they had been provided at the time of the interview, when the interviewee had had the protected time and the safe space to respond spontaneously. No personal names or identifying remarks or subjects mentioned in the audio recordings were written in the transcripts. Participant quotations were used to illustrate the themes, but in line with assurances of anonymity given to interviewees, and no personal identifiers were used. Transcribed interviews were checked against the audio recordings for accuracy before the

data were subject to thematic analysis, guided by Braun and Clarke's [47] 6-step framework, with frequent returning to the data to sense check as we progressed through the framework. Codes were shared with the coauthors (AB, JC, and PM), and a codebook was generated. JM was aware that the thematic analysis and interpretation were influenced by previous experience and assumptions, reflexivity, subjectivity, bias, and emotion. Therefore, in line with Braun and Clarke's advice, a self-reflexive approach was taken throughout the research, including creating a rolling memo of reflections from the time of the interviews until the end of the study. These notes were shared with the coauthors during the stages of thematic analysis.

Results

A total of 14 participants, 10 women and 4 men, with 9 to 34 years of experience as a teacher, completed the interview.

Overview of Findings

Teachers were unanimous in both their recognition of workplace stress and their sincere commitment to teaching. Participants revealed the physical and cultural workplace context to be one of relentless all-direction exposure, multifaceted delivery objectives, and fear of peer opinion and one in which their well-being was less important to policy makers and school heads than that of students. They described the constraining effects of the school context, which significantly affected teachers' ability to look after their well-being. Stress management strategies, except for mindfulness in one school, were self-taught and included immediate real-time or deferred in-school practices, management practices external to the workplace as well as cumulative learning approaches designed to build coping strategies. Many teachers experienced burnout before finding successful stress self-management strategies, which included changing schools.

The pace of work meant that real-time use of health apps by teachers was rare. Outside of school, apps were used for symptom control (eg, insomnia), sometimes automatically (embedded usage that had become second nature, such as diary app entries) and frequently for relaxation and the creation of sustainable habits (such as mindfulness). Wearables were accessed by some teachers in the workplace and often used for quantification, occasionally for prompts (eg, to move or to exercise) or rewards, but wearers' enthusiasm was rarely shared with colleagues.

There was evidence of a shifting organizational culture with the introduction of the concept of teachers' well-being, and most staff members were both very positive about their school leadership and potentially open to leadership recommendations of supportive technology.

The findings reflect 2 overarching themes of understanding how the education workplace context constrains stress self-management and teachers' use of digital health technology. Duration of teaching experience is shown in years after transcript quotes alongside the participant's (P) number and sex (M or F).

Organizing Theme A: Educational Practice Constrains Teachers' Well-Being and Self-Management of Stress

Theme 1: Contextual Constraints on the Use of Technology for Stress Self-Management

Teachers described a context of work that demanded constant exposure on the front line, with their activity simultaneously driven by timetables and interrupted from all directions, which was nonstop in nature.

Theme 1.1: Relentless and Often Simultaneous Demands Dominate the Day

The relentless nature of the teacher's role was illustrated by descriptions of not only delivering back-to-back lessons to class audience of up to 30 (or more) students, all with differing needs and abilities, but also having to manage junior staff members or act on their requests and respond to changeable situations. Not having any time to themselves during the day was frequently mentioned:

> *So as well as being a teacher I'm also head of house. So I've got the additional responsibility of that... it's unpredictable (in) nature because you've got no idea what's going to happen but then obviously the time constraints of when someone has an issue and you're teaching full day it's the time you then have to actually deal with that situation.* [Quote 1, P9, F, 9 years]

> *I don't think a lot of people who aren't in teaching totally get the fact. Oh, you get all these holidays and stuff. It's like, you try teaching. You try presenting to groups of kids for five hours a day. And then getting bombarded with things that you can't just not switch off from.* [Quote 2, P4, F, 16 years]

Theme 1.2: Social and Physical Exposure of the Teacher's Role

The nonstop visibility of the job was an additional issue, with analogies made to acting and becoming *talked out*, having had to present and talk to people all day:

> *I mean I think that is one of the most exhausting things about teaching is that you are performing a role, constantly...And even if you are incredibly stressed, you have to perform a role anyway. And the responsibility and accountability on your shoulders combined with the fact that you are acting most of the day is exhausting.* [Quote 3, P4, F, 16 years]

There was a tangible sense of the social and physical environment dominating their sense of agency and their dignity, with no opportunity for privacy, compromising their ability to self-manage the frequent stressors.

Theme 2: Organizational Culture Influences Self-Management of Stress

Participants went on to describe organizational and cultural reasons for the lack of opportunity to manage stress. The school social environment was described as one in which many teachers still feared admitting the need for help, and for some, this had been exacerbated by the perception that student welfare was

promoted over workforce well-being because of government policy.

Theme 2.1: Stigma of Not Coping Deters Help Seeking

Numerous fears were expressed on how help seeking would be perceived, with concerns including loss of job or demotion, negative peer opinion, and sense of failure. This sometimes extended to the staff not sharing helpful digital tools with peers despite occasionally recommending them to students:

> *No, [I haven't recommended Mindshift to colleagues]. I think we're quite secretive by nature. I think it's such a stigma. As teachers I think we've all got a character trait that we like to do well and I think if you admit you're stressed it's like admitting you're failing.* [Quote 4, P10, F, 22 years]

> *I say it's for research purposes, but actually it's for me as well that I access these things. I use Pacifica.* [Quote 5, P12, F, 32 years]

Many participants' concerns reflected a fear of peer perception that they might not be coping well, highlighting a culture where weaknesses were seen as having their origin in the individual rather than in the system.

Theme 2.2: Teachers' Well-Being Secondary to Student Well-Being

Added to this fear of an unsympathetic social culture was the inference, from recent national policy, that the focus was solely on prioritizing student well-being, as well as the growing emphasis in the education sector on teachers meeting the mental health needs of students with little provision for their own:

> *You've really got to look after yourself because nobody else is going to look after you here. They're not because they can't...nobody's got time to look after their staff.* [Quote 6, P7, F, 16 years]

> *So you've got stressed teachers trying to un-stress children which doesn't make much sense.* [Quote 7, P9, F, 9 years]

Despite the rhetoric on well-being, the message that it is "ok to not be ok" reported by a few participants has not yet been normalized.

Theme 3: Feeling Isolated in School, but Indications That School Culture is Changing

Sometimes participants spoke about feeling isolated, but this was often implied through descriptions of trying to manage stressful times on their own. Others, notably within the smaller setting of subject departments, had been very upfront about struggles and had shared simple advice with colleagues, such as walking more slowly between lessons, or had already suggested technology solutions to stressful scenarios.

Theme 3.1: Individual Scenarios of Isolation

Sometimes, teachers described scenarios where stress had been overwhelming and the early warning symptoms of insomnia or nervous tics were ignored. These participants were frequently conscious of work dominating their lives, but in their isolation, they saw the ability to cope as an individual responsibility. Subsequently, their well-being suffered with occasionally devastating consequences such as nervous or marriage breakdown:

> *I went down to my staff base. I moved myself away from the door so nobody could look in, went to one of the last compartments, and I cried and I cried and I cried. There's nobody you can go to. When you're in this environment there's nobody. There's nobody to be able to download to. There's nobody to say are you alright. You're just completely by yourself. So the feeling of isolation...and actually...quite powerless to do things.* [Quote 8, P7, F, 16 years]

> *I used to bring three bottles of water to school and I could tell if I'd had a bad day because...I've had a sip out of each when I got home and I didn't need to go to the loo. And there'd have been no space.* [Quote 9, P10, F, 22 years]

Theme 3.2: Tackling Isolation Within a Local Culture

Descriptions of subject departments were usually very warm, with heads of department describing camaraderie among colleagues, which led to a more supportive culture:

> *Yes, I'm known for emailing [my department] and saying use this app, I find it really useful. And I did some work looking at reducing workload and marking and being really sort of honest with staff and saying I'm really struggling with this. I struggle with workload and I want to try and make it better. So I share quite a lot.* [Quote 10, P6, F, 20 years]

> *I set up a support programme for [my colleague], based on my experience. What would really suit you? What are you happy to do? In terms of the aspects that I could control, I came to an agreement and a balance with her, just making sure to check up on her.* [Quote 11, P14, F, 15 years]

A majority of participants believed that the wider leadership sincerely wanted to support staff well-being, some referencing an annual staff well-being survey as well as some responses that had been acted upon.

Theme 3.3: Raising the Profile of Teacher Well-Being

All the staff members interviewed were or had been in a position of leadership themselves, and some were involved in their schools' plans to raise the profile and provision of well-being for the staff. Participants were wary about gimmicks; some had noticed mental health posters appearing in their school but without an accompanying narrative. However, the suggestion that leadership could proactively recommend solutions raised genuine interest:

> *My job at the minute in terms of staff wellbeing and staff development is partly because there are changes that I wish to make because I've been here for a long time...I've listened to people, I've got a sense of what people say and wish to do and actually avoiding helplessness, this is something I can actually do and make that happen.* [Quote 12, P11, F, 22 years]

> *No, I would love (the Senior Leadership Team to recommend a wearable that is really good for*

> *managing stress). I like stuff like that. No, that is fine.* [Quote 13, P3, F, 28 years]

Despite expressions of a constricting culture in which stress management was seen as an individual and not an institutional responsibility, there was still a strong sense of trust in the leadership. Most of the staff members had worked at more than one school, and many had changed schools to find a culture in which they felt more supported or valued.

Organizing Theme B: Some Digital Stress Technologies Can Be Used and Could Be Expanded

Theme 4: Stress Reduction Technologies Are Considered Useful by Teachers

Across the participants interviewed, 12 out of 14 participants already used digital tools for stress management or well-being, with 17 apps, 5 different brands of wearables, and 1 desktop program named.

Theme 4.1: Both the Form of Technology and Physical Workspace Influence Its Use

Smartphones were taken to school by teachers, but time, signal, and privacy constraints meant that phones remained in the teacher's bag or coat pocket; just one teacher accessed a calming app in the privacy of her office during the day, and another accessed the Headspace app via her iPad. Over half of those interviewed used wearable technology (eg, Fitbit and smart watch) and real-time data to manage their stress and well-being:

> *I used [Headspace] when I was in the bath last night...just focusing on breathing and meditation. And I did it the previous night in bed, I just did like five minutes and even just after five minutes I felt quite relaxed and I thought wow, I should just do five minutes a day because it's nothing really...I think if I didn't have that app I wouldn't do it and it now sends me reminders.* [Quote 14, P2, M, 17 years]

> *Anyway, back to the Fitbit. I just love all the features of it. It's got relax as well. Sometimes if I wake up in the middle of the night, I put the relax on. Quick, two-minute session to find a bit of calm.* [Quote 15, P3, F, 28 years]

Theme 4.2: Teachers Are Open to Schools Promoting Destressing Technologies

When asked about their response to the school leadership promoting a digital health innovation for well-being, interviewees voiced near unanimous enthusiasm and a belief that this would only happen if due diligence had been already done to conclude that such a tool was tried and trusted. They might not feel it was for them, but it could be for a colleague:

> *I'd look at that [stress app]. I'd look at anything because it doesn't have to help me does it? It becomes part of a toolkit of stuff that you signpost to people...This is stuff I'm prepared to recommend. Actually this becomes our provision. This becomes what we do.* [Quote 16, P11, F, 22 years]

Theme 5: Sequelae of Stress Targeted by a Technology Intervention Informs Potential Choices

In terms of meeting their own needs, teachers managed stress using strategies that included the setting of boundaries, focusing on holistic well-being, keeping a diary, or intentional reflection on experiences. Sequelae of stress, such as elevated heart rate, shallow breathing, or racing thoughts, were often the target of conscious self-management strategies, including via digital health tools. Mindfulness was the most frequently cited stress reduction technique, both with and without digital support. Those who had wearable technology all expressed an interest in observing their patterns of behavior and activity, although they did not always welcome the attendants' prompts. One interviewee had collated data from the wearable to successfully demonstrate to a line manager the need to change her schedule because of the potential exacerbation of a health condition. Many wearers did not share their data or advocate for their wearable despite liking the personal feedback. However, they did appreciate the data captured by technology tools, which was frequently mentioned as valuable and confirmatory.

Theme 5.1: Monitoring Behavior or Mood Patterns to Aid Reflection

The data gathered were used to aid reflection on reactions or responses to stress, and participants expressed the value of seeing patterns of behavior through tracking and mapping:

> *So it's about patterns...My wonderful days. It's a diary [app]. And I am supposed to write down every day what I am feeling but it tends to be just what I have done that day whether it's longer or shorter or...I can look back at last year. You can see how far you've come - oh my goodness that was how I was feeling this time last year!* [Quote 17, P1, M, 34 years]

> *The only thing I've started, and I'm wearing it at the moment, is a heart rate monitor. That's more for when I exercise, but I've been conscious that it comes up with a sleep pattern. Every morning, it's less than the desired amount of sleep achieved, poor quality of sleep. So, it's really interesting to see that, on a regular basis, I'm not getting the right amount of sleep.* [Quote 18, P13, M, 11 years]

Theme 5.2: Use of Technology to Directly Self-Manage Stress Sequelae

Participants used apps to access stress management techniques at home, such as dealing with anxiety or relaxing through mindfulness or meditation. Wearables were able to afford real-time intervention or information, such as helping them get back to sleep at night or bringing down their heart rate through managed breathing during a class in school:

> *And again, often if I feel myself getting a bit stressed [in class] I will look at my Fitbit and if I can see that my heartrate is above 80, which for me is very high, again, I will either do a breathing exercise or I'll just do something to change what I've got myself into and then just check again in a few minutes.* [Quote 19, P6, F, 20 years]

> *Currently, I like Insight Timer. I like it because I can adapt it to whatever I'm doing...At the end of the day...if I'm wound up about something, there is [a programme I can chose for example] for resentment or anxiety. It's very, very adaptable.* [Quote 20, P10, F, 22 years]

Theme 5.3: Use of Technology for More Holistic Self-Care

Some teachers described indirect ways of reducing stress, managing their time better to impact workload or build self-efficacy. An example of this was seen through the use of the research tool TeacherTapp. This app has been designed to capture data from teachers at the end of their working day by asking them 3 research questions. By taking part, teachers themselves can review the findings and read associated articles:

> *It [TeacherTapp]) also links each day to some kind of educational research...and it also tells me how long it takes to read...it kind of makes me feel more professional and that kind of feeds into the stress and the workload thing doesn't it. So although it's not a wellbeing tool I think it does work as one.* [Quote 21, P6, F, 20 years]

Discussion

Principal Findings

This study used a qualitative approach across several real-world settings to develop an understanding of the contextual influences on English midcareer high school teachers' self-management of workplace stress and of their use of digital health tools for well-being. Consequently, we explored each in the following sections.

Educational Practice Constrains Teachers' Well-Being and Self-Management of Stress

The findings from the data analysis formed a detailed narrative of how self-management of workplace stress by senior teachers is subject to constraining contextual and organizational factors. Although the concept of constraints has often been cited in teaching literature [52], with workload the most cited cause of stress [53,54], details of how contextual constraints affect staff well-being or how interventions can help with stress [13] are sparse. The first of these constraints identified by participants was of relentless demands, which have been described elsewhere [55,56]. Yet, when combined with a second constraint of lack of privacy—the ability to withdraw from the front line or absence of self-care time known as *local privacy*—this absence of refuge results in inhibition, stifling of expression, or no opportunity to destress [57]. Some schools no longer provided a staff room for teachers to retreat to, yet we know from Maslach's [58] work that stress leading to burnout is most frequent among those whose work involves intense ongoing involvement with other people, such as doctors and teachers.

Stigma of help seeking was a third constraining factor. This related to teachers' fear of being perceived to be of *inferior ability* if they ask for help [56], as well as the stigma that continues to exist around seeking mental health support. There was a notable exception described by one participant who had a mental illness diagnosis and experienced excellent support and understanding from both the school leadership and students. However, the findings indicate that the cultural normalization of well-being support has not been achieved.

It was more often seen as ironic that the recent national policy's emphasis on mental health in schools was almost exclusively aimed at students and not the education workforce [59], despite evidence of teachers' concerns that if their own emotional needs are neglected they cannot meet those of the students [31,60]. These confidentiality and peer concerns mirror the opinion of health care professionals [61,62], and anonymity is still cited in research as a strong benefit to online mental health support forums [63]. Using Web-based, evidence-based therapy or psychoeducation has also been shown to reduce negative perceptions of mental illness, so there is potentially a double beneficial effect with treatment and attitude [64-66]. These factors of intensity, exposure, and apparent sidelining need to inform stress support technology design and implementation for teachers. They should also inform education policy leaders' planning for whole-school mental health initiatives.

"Feeling lonely in a crowded room" was how one teacher described one symptom of burnout, and it was poignant how symptoms such as isolation, dehydration, hunger, and insomnia were frequently described alongside the assumption that it was purely down to the individual to manage the causes and manifestations of stress. This individualized responsibility can place a profound burden on a teacher as such an approach fails to acknowledge the complexity of the origins of stress [65]. Aspirations of the school leadership to raise the profile of teacher well-being could be harnessed through an organizational approach to address these common causes. The trust expressed in the leadership, despite the understanding that the leadership's hands are sometimes tied by external factors such as imposed student number increases and testing, appears to be a valuable opportunity to acknowledge and consider interventions that could ameliorate stress indicators at an organizational level.

Some Digital Stress Technologies Can Be Used and Could Be Expanded

Our second organizing theme extends previous research on the use of health technology for workplace well-being. Smartphones (and hence real-time health apps) were rarely accessed during the working day by teachers. Among the 23 mental eHealth (Web-based) at-work studies reviewed by Stratton [35] in 2017, only 3 studies seemed to utilize mobile apps, indicating that smartphones were used in a small minority of studies. Just one eHealth study was conducted among teachers, using a Web-based program. No mention was made of when the program was accessed during the day or whether it was possible to access the program via an app [41]. This study's findings do not rule out teachers' stress management being assisted through an app, but we can for the first time infer that apps would not be suitable for just-in-time stress relief, certainly not in the presently constructed school context.

However, the participants were nearly unanimous in their willingness to give a particular stress-relieving or prevention technology a try should the school ever promote such a technology. This could partly be a reflection of the high rate of

participant use of well-being technology, which was already present in our sample (12 out of 14 participants), but it was also a clear expression of their trust in the school leadership. Most participants believed that the school leadership would only endorse a stress-relieving technology that could be valuable and reliable. This could reflect the fact that the staff members worked at schools that acknowledged staff well-being or that the staff members had chosen to work in a school where they felt valued, which engendered their trust in the leadership.

Current use of technology by teachers demonstrated the potential for combining self-management of stress and organizational promotion of well-being. However, the proportion of teachers who did (12/14, 86%) cannot be extrapolated to the whole population as they self-selected for the interview and were not a random sample. Of Nunes' description of 5 modus operandi (MO) of technology for self-care, 2 of the 5 were found to be individually oriented: (1) fostering reflection and (2) suggesting treatment. The 3 other MO's were communal: (3) sharing care activity, (4) enhancing collaboration, and (5) peer-to-peer support [67]. This study showed how these MO have been used in the education workplace providing valuable indicators both for potential support tools and delivery of an organizational intervention.

Among participants, methods that fostered reflection, MO (1), were the most widely used both with and without technology. In fact, 2 participants revealed how technology had become embedded, realizing during the interview that it was their apps that aided reflection. Of the 12 participants who reported using technology to support their well-being, 8 participants reported having a wearable band or watch. The real-time feedback and actionable insights of the wearables (2) were felt to be useful and empowering, enabling positive decisions to be taken to manage stress indicators. These included undertaking a breathing exercise to lower an elevated heart rate or planning more exercise. Feedback on and insight into sleep quality were mentioned by several people as provoking a change in behavior. These examples correlate with the approach of empowerment through meaningful data, facilitating participants to make changes in a way that also respects their autonomy as described by Tengland [68]. This author makes the important point that empowerment can be seen as ethically preferable to behavior change, as the latter is often concealed from the users, whereas empowerment directly informs and assists individuals, enabling better choices to be made consciously by them. Given the value placed on trust in the leadership by teachers, any organizationally directed intervention would have to be empowering and transparent.

Most teachers did not promote their wearable to colleagues, but 2 participants did mention enjoying comparing activity with other teachers as per the peer-to-peer MO (5) noted above. A recent systematic review of workplace studies utilizing mobile health, including wearable interventions to promote physical activity, has shown reasonable evidence for workplace eHealth in a workplace context as "a feasible, acceptable and effective tool" [37]. This study's findings reveal the contextual value placed on wearables by some teachers and suggest an acceptability of this modality.

Many of the apps used facilitated immediate therapeutic access, most commonly to mindfulness, breathing, or meditative tools, combining Nunes' (1) and (2) MO descriptions. Only one teacher claimed to use Web-based CBT, but some of the apps mentioned by the staff used CBT principles. The staff also engaged with both mindfulness and social support (eg, fit2teach) technologies. All 3 strategies aligned with the evidence for stress reduction efficacy among teachers [21], and these show self-management choices that could be of value to other senior teachers. Only the Web-based CBT resulted from a professional referral. One participant described how using the TeacherTapp app, which has been designed for education research [69], not only enhanced collaboration (the communal aspect of collaboration from Nunes' description of MO (4) of technology for self-care) but also made her feel more professional, providing an indirect boost to her well-being. This participant also noted how forums on social media enabled peer-to-peer support. This was for professional sharing rather than direct stress management, but it still had a positive effect in terms of reducing the workload.

Most tools were adopted serendipitously, discovered through advertising, friends and family, or social media, or just occasionally through peer recommendation. Both the inclination of the staff and availability of technology suggest opportunities for technology-enhanced stress self-management that both reflect current strategies and could possibly work around the described constraints.

The UK government's 2017 Farmer-Stevenson report included recommending public sector employers to "Ensure provision of tailored in-house mental health support and signposting to clinical help" [70]. It is hoped that the findings of this study will go some way to enabling that *tailored* support.

Strengths and Limitations

This was a small-scale study, and the participant sample cannot be claimed to be representative of all midcareer high school teachers. In particular, as the sample was self-selecting, it is likely that there were a higher proportion of staff members who had experienced crises and time out of teaching than the average population. Beyond the confirmation of the school's recognition of staff well-being, this study did not investigate the influence of the school's organizational or leadership attitude on how teachers managed their stress. The participants knew in advance that they were going to be asked about their use of health technology; therefore, although 2 participants did not use any, the proportion of those who did cannot be generalized to the whole teaching population. In addition, this study was limited by a lack of ethnic or racial diversity.

All the staff members were interviewed during term time after at least five weeks back at school after the summer holidays. It was intentional that participants were interviewed while in the thick of school activity, and all of them were very much immersed in the demands of school life. Owing to their busy schedules, we are immensely grateful to the participants who gave their time, courageously talked about stressful situations, and who without exception demonstrated commitment and determination. Their pride in their students was tangible.

As a former clinician, JM was trained to look for symptoms and think diagnostically. Inevitably, although this can aid insight, it could mean a more medical bias to the analysis. However, the coauthors were from nonclinical backgrounds, and all the findings were reviewed with them.

Conclusions

The aim of this study was to elicit the details of the contextual influences on midcareer high school teachers' management of stress. This study's findings are consistent with previous research among education and health care professionals on how context shapes their stress management. Distinctively, we described both the cultural and physical contextual reasons, including endless demands, lack of local privacy, stigma, inferiority, and individualism, which are constraining teachers' workplace stress management. This study also indicated important signs or correlates of stress for teachers, including dehydration, feelings of isolation, and insomnia.

This research has provided distinct insights into how DHIs are being used for stress management and noted that mobile apps have the most potential for stress support outside the workplace, although wearable technologies can provide easy access to data within the workplace.

The important directions indicated for future teacher studies are to explore contextual changes to improve stress management, examine the correlates of stress that could be addressed by technology support, identify how and why stress symptoms can be helped using digital tools, and determine how such tools can be designed to provide enhanced, tailored support.

Acknowledgments

This research received funding from Engineering and Physical Sciences Research Council Doctoral Training Partnership awarded to UCL EP/R513143/1.

Authors' Contributions

All authors reviewed and provided feedback on the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The semistructured qualitative interview guide.
[DOCX File , 18 KB - mental_v7i4e15416_app1.docx]

References

1. Johnson S, Cooper C, Cartwright S, Donald I, Taylor P, Millet C. The experience of work - related stress across occupations. J Manag Psychol 2005;20(2):178-187. [doi: 10.1108/02683940510579803]
2. Schneiderman N, Ironson G, Siegel SD. Stress and health: psychological, behavioral, and biological determinants. Annu Rev Clin Psychol 2005;1:607-628 [FREE Full text] [doi: 10.1146/annurev.clinpsy.1.102803.144141] [Medline: 17716101]
3. Kyriacou C, Sutcliffe J. Teacher stress: prevalence, sources, and symptoms. Br J Educ Psychol 1978 Jun;48(2):159-167. [doi: 10.1111/j.2044-8279.1978.tb02381.x] [Medline: 687523]
4. Cooper CL. Life at the chalkface--identifying and measuring teacher stress. Br J Educ Psychol 1995 Mar;65 (Pt 1):69-71. [doi: 10.1111/j.2044-8279.1995.tb01131.x] [Medline: 7727268]
5. Cooper CL, Kelly M. Occupational stress in head teachers: a national UK study. Br J Educ Psychol 1993 Feb;63 (Pt 1):130-143. [doi: 10.1111/j.2044-8279.1993.tb01046.x] [Medline: 8466831]
6. Travers C, Cooper C. Teachers Under Pressure: Stress In The Teaching Profession. Abingdon: Routledge; 2012.
7. McCarthy CJ, Lambert RG, Lineback S, Fitchett P, Baddouh PG. Assessing teacher appraisals and stress in the classroom: review of the classroom appraisal of resources and demands. Educ Psychol Rev 2016;28(3):577-603. [doi: 10.1007/s10648-015-9322-6]
8. Reiser JE, Murphy SL, McCarthy CJ. Stress prevention and mindfulness: a psychoeducational and support group for teachers. J Spec Gr Work 2016;41(2):117-139. [doi: 10.1080/01933922.2016.1151470]
9. Camacho DA, Vera E, Scardamalia K, Phalen PL. What are urban teachers thinking and feeling? Psychol Sch 2018;55(9):1133-1150. [doi: 10.1002/pits.22176]
10. Anderson CS. The search for school climate: a review of the research. Rev Educ Res 1982;52(3):368-420 [FREE Full text] [doi: 10.2307/1170423]
11. Loukas A. National Association of Elementary School Principals: NAESP. 2017. What Is School Climate? Leaders Compass 2007;5(1) URL: https://www.naesp.org/sites/default/files/resources/2/Leadership_Compass/2007/LC2007v5n1a4.pdf [accessed 2019-07-08]
12. Brown Z, Uehara D. Education Resources Information Center. 1999. Coping with Teacher Stress: A Research Synthesis for Pacific Educators. Research Series URL: https://tinyurl.com/yy7qoz49 [accessed 2019-07-08]

13. Jeter L. Education Resources Information Center. 2012. Coping Strategies Title I Teachers Use to Manage Burnout and Stress: A Multisite Case Study URL: https://eric.ed.gov/?id=ED554626 [accessed 2019-07-08]
14. Wilkerson K. An examination of burnout among school counselors guided by stress-strain-coping theory. J Couns Dev 2009;87(4):428-437. [doi: 10.1002/j.1556-6678.2009.tb00127.x]
15. Sharma M, Rush SE. Mindfulness-based stress reduction as a stress management intervention for healthy individuals: a systematic review. J Evid Based Complementary Altern Med 2014 Oct;19(4):271-286. [doi: 10.1177/2156587214543143] [Medline: 25053754]
16. Khoury B, Sharma M, Rush SE, Fournier C. Mindfulness-based stress reduction for healthy individuals: a meta-analysis. J Psychosom Res 2015 Jun;78(6):519-528. [doi: 10.1016/j.jpsychores.2015.03.009] [Medline: 25818837]
17. Greenberg M, Brown J, Abenavoli R. Robert Wood Johnson Foundation. 2016. Teacher Stress and Health: Effects on Teachers, Students, and Schools URL: https://www.rwjf.org/en/library/research/2016/07/teacher-stress-and-health.html [accessed 2019-06-05] [WebCite Cache ID 78u9fQpcQ]
18. Ouellette RR, Frazier SL, Shernoff ES, Cappella E, Mehta TG, Maríñez-Lora A, et al. Teacher job stress and satisfaction in urban schools: disentangling individual-, classroom-, and organizational-level influences. Behav Ther 2018 Jul;49(4):494-508 [FREE Full text] [doi: 10.1016/j.beth.2017.11.011] [Medline: 29937253]
19. Naghieh A, Montgomery P, Bonell CP, Thompson M, Aber JL. Organisational interventions for improving wellbeing and reducing work-related stress in teachers. Cochrane Database Syst Rev 2015 Apr 8(4):CD010306. [doi: 10.1002/14651858.CD010306.pub2] [Medline: 25851427]
20. Mansfield CF, Beltman S, Broadley T, Weatherby-Fell N. Building resilience in teacher education: an evidenced informed framework. Teach Teach Educ 2016 Feb;54:77-87. [doi: 10.1016/j.tate.2015.11.016]
21. Iancu AE, Rusu A, Măroiu C, Păcurar R, Maricu oiu LP. The effectiveness of interventions aimed at reducing teacher burnout: a meta-analysis. Educ Psychol Rev 2018;30(2):373-396. [doi: 10.1007/s10648-017-9420-8]
22. Roeser RW, Schonert-Reichl KA, Jha A, Cullen M, Wallace L, Wilensky R, et al. Mindfulness training and reductions in teacher stress and burnout: results from two randomized, waitlist-control field trials. J Educ Psychol 2013;105(3):787-804. [doi: 10.1037/a0032093]
23. Emery DW. University of Missouri-St Louis. 2011. Crisis in Education: A Call to ACT URL: https://irl.umsl.edu/dissertation/426/ [accessed 2019-07-08]
24. Unterbrink T, Pfeifer R, Krippeit L, Zimmermann L, Rose U, Joos A, et al. Burnout and effort-reward imbalance improvement for teachers by a manual-based group program. Int Arch Occup Environ Health 2012 Aug;85(6):667-674. [doi: 10.1007/s00420-011-0712-x] [Medline: 22038086]
25. Patton G, Bond L, Butler H, Glover S. Changing schools, changing health? Design and implementation of the Gatehouse Project. J Adolesc Health 2003;33(4):231-239. [doi: 10.1016/s1054-139x(03)00204-0]
26. Wyn J, Cahill H, Holdsworth R, Rowling L, Carson S. MindMatters, a whole-school approach promoting mental health and wellbeing. Aust N Z J Psychiatry 2000 Aug;34(4):594-601. [doi: 10.1080/j.1440-1614.2000.00748.x] [Medline: 10954390]
27. Willis A, Hyde M, Black A. Juggling with both hands tied behind my back: teachers' views and experiences of the tensions between student well-being concerns and academic performance improvement agendas. Am Educ Res J 2019;56(6):2644-2673. [doi: 10.3102/0002831219849877]
28. Anderson M, Werner-Seidler A, King C, Gayed A, Harvey SB, O'Dea B. Mental health training programs for secondary school teachers: a systematic review. School Ment Health 2018;11(3):489-508. [doi: 10.1007/s12310-018-9291-2]
29. Kidger J, Brockman R, Tilling K, Campbell R, Ford T, Araya R, et al. Teachers' wellbeing and depressive symptoms, and associated risk factors: A large cross sectional study in English secondary schools. J Affect Disord 2016 Mar 1;192:76-82 [FREE Full text] [doi: 10.1016/j.jad.2015.11.054] [Medline: 26707351]
30. Jennings PA, Greenberg MT. The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. Rev Educ Res 2009;79(1):491-525. [doi: 10.3102/0034654308325693]
31. Sisask M, Värnik P, Värnik A, Apter A, Balazs J, Balint M, et al. Teacher satisfaction with school and psychological well-being affects their readiness to help children with mental health problems. Health Educ J 2014;73(4):382-393. [doi: 10.1177/0017896913485742]
32. Rowe F, Stewart D, Patterson C. Promoting school connectedness through whole school approaches. Health Educ 2007;107(6):524-542. [doi: 10.1108/09654280710827920]
33. Kuster AT, Dalsbø TK, Thanh BY, Agarwal A, Durand-Moreau QV, Kirkehei I. Computer-based versus in-person interventions for preventing and reducing stress in workers. Cochrane Database Syst Rev 2017 Aug 30;8:CD011899 [FREE Full text] [doi: 10.1002/14651858.CD011899.pub2] [Medline: 28853146]
34. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. J Med Internet Res 2017 Jul 26;19(7):e271 [FREE Full text] [doi: 10.2196/jmir.7583] [Medline: 28747293]
35. Stratton E, Lampit A, Choi I, Calvo RA, Harvey SB, Glozier N. Effectiveness of eHealth interventions for reducing mental health conditions in employees: a systematic review and meta-analysis. PLoS One 2017;12(12):e0189904 [FREE Full text] [doi: 10.1371/journal.pone.0189904] [Medline: 29267334]

36. Tucker SJ, Lanningham-Foster LM, Murphy JN, Thompson WG, Weymiller AJ, Lohse C, et al. Effects of a worksite physical activity intervention for hospital nurses who are working mothers. AAOHN J 2011 Sep;59(9):377-386. [doi: 10.3928/08910162-20110825-01] [Medline: 21877670]
37. Buckingham SA, Williams AJ, Morrissey K, Price L, Harrison J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review. Digit Health 2019;5:2055207619839883 [FREE Full text] [doi: 10.1177/2055207619839883] [Medline: 30944728]
38. Zenonos A, Khan A, Kalogridis G, Vatsikas S, Lewis T, Sooriyabandara M. HealthyOffice: Mood Recognition At Work Using Smartphones and Wearable Sensors. In: 2016 IEEE International Conference on Pervasive Computing and Communication Workshops. 2016 Presented at: PerComW'16; March 14-18, 2016; Sydney, NSW, Australia p. 14-18. [doi: 10.1109/percomw.2016.7457166]
39. Can YS, Arnrich B, Ersoy C. Stress detection in daily life scenarios using smart phones and wearable sensors: a survey. J Biomed Inform 2019 Apr;92:103139. [doi: 10.1016/j.jbi.2019.103139] [Medline: 30825538]
40. Arza A, Garzón-Rey JM, Lázaro J, Gil E, Lopez-Anton R, de la Camara C, et al. Measuring acute stress response through physiological signals: towards a quantitative assessment of stress. Med Biol Eng Comput 2019 Jan;57(1):271-287. [doi: 10.1007/s11517-018-1879-z] [Medline: 30094756]
41. Ebert DD, Lehr D, Boß L, Riper H, Cuijpers P, Andersson G, et al. Efficacy of an internet-based problem-solving training for teachers: results of a randomized controlled trial. Scand J Work Environ Health 2014 Nov;40(6):582-596 [FREE Full text] [doi: 10.5271/sjweh.3449] [Medline: 25121986]
42. Anderson L, Olsen B. Investigating early career urban teachers' perspectives on and experiences in professional development. J Teach Educ 2006;57(4):359-377. [doi: 10.1177/0022487106291565]
43. Freund AM, Baltes PB. Life-management strategies of selection, optimization, and compensation: measurement by self-report and construct validity. J Pers Soc Psychol 2002 Apr;82(4):642-662. [Medline: 11999929]
44. Bradley G. Job tenure as a moderator of stressor–strain relations: a comparison of experienced and new-start teachers. Work Stress 2007 Jan;21(1):48-64. [doi: 10.1080/02678370701264685]
45. Philipp A, Kunter M. How do teachers spend their time? A study on teachers' strategies of selection, optimisation, and compensation over their career cycle. Teach Teach Educ 2013;35:1-12. [doi: 10.1016/j.tate.2013.04.014]
46. Shain M, Kramer DM. Health promotion in the workplace: framing the concept; reviewing the evidence. Occup Environ Med 2004 Jul;61(7):643-8, 585 [FREE Full text] [doi: 10.1136/oem.2004.013193] [Medline: 15208383]
47. Clarke V, Braun V. Successful Qualitative Research: A Practical Guide For Beginners. California: Sage Publications Ltd; 2013.
48. Education Support Partnership. Stress Test URL: https://tinyurl.com/y6hhqnfh [accessed 2019-07-08]
49. Smith A, Brice C, Collins A, Matthews V, McNamara R. The Scale of Occupational Stress: A Further Analysis of the Impact of Demographic Factors and Type of Job. Bootle, Merseyside: Health And Safety Executive (HSE); 2000.
50. Health and Safety Executive. Managing The Causes Of Work-related Stress: A Step-by-Step Approach Using The Management Standards. Norwich, UK: Health and Safety Executive; 2007.
51. Baker SE, Edwards R. National Centre for Research Methods. Southampton, UK; 2012. How Many Qualitative Interviews is Enough? Expert Voices and Early Career Reflections on Sampling and Cases in Qualitative Research URL: https://tinyurl.com/nexe5yy [accessed 2019-07-08]
52. Payne R, Fletcher B. Job demands, supports, and constraints as predictors of psychological strain among schoolteachers. J Vocat Behav 1983;22(2):136-147. [doi: 10.1016/0001-8791(83)90023-4]
53. Alarcon GM. A meta-analysis of burnout with job demands, resources, and attitudes. J Vocat Behav 2011;79(2):549-562. [doi: 10.1016/j.jvb.2011.03.007]
54. Avanzi L, Fraccaroli F, Castelli L, Marcionetti J, Crescentini A, Balducci C, et al. How to mobilize social support against workload and burnout: The role of organizational identification. Teach Teach Educ 2018;69:154-167. [doi: 10.1016/j.tate.2017.10.001]
55. Smith M, Bourke S. Teacher stress: Examining a model based on context, workload, and satisfaction. Teach Teach Educ 1992;8(1):31-46. [doi: 10.1016/0742-051X(92)90038-5]
56. Butler R. Teachers' achievement goal orientations and associations with teachers' help seeking: examination of a novel approach to teacher motivation. J Educ Psychol 2007;99(2):241-252. [doi: 10.1037/0022-0663.99.2.241]
57. Palm E. Privacy expectations at work—what is reasonable and why? Ethic Theory Moral Prac 2009;12(2):201-215. [doi: 10.1007/s10677-008-9129-3]
58. Maslach C, Schaufeli WB, Leiter MP. Job burnout. Annu Rev Psychol 2001;52:397-422. [doi: 10.1146/annurev.psych.52.1.397] [Medline: 11148311]
59. Department of Health and Social Care. The Government of UK. Secondary School Staff Get Mental Health 'First Aid' Training URL: https://tinyurl.com/yxe3f9oo [accessed 2019-07-08]
60. Kidger J, Gunnell D, Biddle L, Campbell R, Donovan J. Part and parcel of teaching? Secondary school staff's views on supporting student emotional health and well‐being. Br Educ Res J 2009;36(6):919-935 [FREE Full text] [doi: 10.1080/01411920903249308]

61. Montgomery A. The inevitability of physician burnout: Implications for interventions. Burn Res 2014;1(1):50-56 [FREE Full text] [doi: 10.1016/j.burn.2014.04.002]
62. Spiers J, Buszewicz M, Chew-Graham CA, Gerada C, Kessler D, Leggett N, et al. Barriers, facilitators, and survival strategies for GPs seeking treatment for distress: a qualitative study. Br J Gen Pract 2017 Oct;67(663):e700-e708 [FREE Full text] [doi: 10.3399/bjgp17X692573] [Medline: 28893766]
63. Smith-Merry J, Goggin G, Campbell A, McKenzie K, Ridout B, Baylosis C. Social connection and online engagement: insights from interviews with users of a mental health online forum. JMIR Ment Health 2019 Mar 26;6(3):e11084 [FREE Full text] [doi: 10.2196/11084] [Medline: 30912760]
64. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br J Psychiatry 2004 Oct;185:342-349. [doi: 10.1192/bjp.185.4.342] [Medline: 15458995]
65. Taylor-Rodgers E, Batterham PJ. Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: randomised controlled trial. J Affect Disord 2014 Oct;168:65-71. [doi: 10.1016/j.jad.2014.06.047] [Medline: 25038293]
66. Kelly P, Colquhoun D. The professionalization of stress management: health and well-being as a professional duty of care? Crit Pub Health 2005;15(2):135-145 [FREE Full text] [doi: 10.1080/09581590500144942]
67. Nunes F, Verdezoto N, Fitzpatrick G, Kyng M, Grönvall E, Storni C. Self-care technologies in HCI. ACM Trans Comput-Hum Interact 2015;22(6):1-45. [doi: 10.1145/2803173]
68. Tengland PA. Behavior change or empowerment: on the ethics of health-promotion strategies. Pub Health Ethics 2012;5(2):140-153. [doi: 10.1093/phe/phs022]
69. McInerney L. Teacher Tapp. Some Tappers Are Worth More Than Others! — How We 'Weight' the Teacher Tapp Sample URL: https://teachertapp.co.uk/some-tappers-worth-more-reweighting-sample/ [accessed 2019-05-22] [WebCite Cache ID 78YZ4PwnE]
70. Farmer P, Stevenson D. The UK Government. 2017. Thriving at Work: A Review of Mental Health and Employers URL: https://tinyurl.com/yy28mfpg [accessed 2019-07-08]

Abbreviations

CBT: cognitive behavioral therapy
DHI: digital health intervention
eHealth: electronic health
MO: modus operandi
PST: problem solving therapy
RCT: randomized controlled trial

Edited by J Torous; submitted 09.07.19; peer-reviewed by S Horvath, J Appalasamy; comments to author 02.09.19; accepted 10.09.19; published 27.04.20.

Please cite as:
Manning JB, Blandford A, Edbrooke-Childs J, Marshall P
How Contextual Constraints Shape Midcareer High School Teachers' Stress Management and Use of Digital Support Tools: Qualitative Study
JMIR Ment Health 2020;7(4):e15416
URL: http://mental.jmir.org/2020/4/e15416/
doi:10.2196/15416
PMID:32338623

Original Paper

Analyzing Trends of Loneliness Through Large-Scale Analysis of Social Media Postings: Observational Study

Keren Mazuz[1], PhD; Elad Yom-Tov[2,3], PhD

[1]Hadassah Academic College Jerusalem, Jerusalem, Israel

[2]Microsoft Research, Herzeliya, Israel

[3]Faculty of Industrial Engineering and Management, Technion, Haifa, Israel

Corresponding Author:
Elad Yom-Tov, PhD
Microsoft Research
13 Shenkar St
Herzeliya, 46733
Israel
Phone: 972 747111359
Email: eladyt@yahoo.com

Abstract

Background: Loneliness has become a public health problem described as an epidemic, and it has been argued that digital behavior such as social media posting affects loneliness.

Objective: The aim of this study is to expand knowledge of the determinants of loneliness by investigating online postings in a social media forum devoted to loneliness. Specifically, this study aims to analyze the temporal trends in loneliness and their associations with topics of interest, especially with those related to mental health determinants.

Methods: We collected a total of 19,668 postings from 11,054 users in the loneliness forum on Reddit. We asked seven crowdsourced workers to imagine themselves as writing 1 of 236 randomly chosen posts and to answer the short-form UCLA Loneliness Scale. After showing that these postings could provide an assessment of loneliness, we built a predictive model for loneliness scores based on the posts' text and applied it to all collected postings. We then analyzed trends in loneliness postings over time and their correlations with other topics of interest related to mental health determinants.

Results: We found that crowdsourced workers can estimate loneliness (interclass correlation=0.19) and that predictive models are correlated with reported loneliness scores (Pearson r=0.38). Our results show that increases in loneliness are strongly associated with postings to a suicidality-related forum (hazard ratio 1.19) and to forums associated with other detrimental behaviors such as depression and illicit drug use. Clustering demonstrates that people who are lonely come from diverse demographics and from a variety of interests.

Conclusions: The results demonstrate that it is possible for unrelated individuals to assess people's social media postings for loneliness. Moreover, our findings show the multidimensional nature of online loneliness and its correlated behaviors. Our study shows the advantages of studying a hard-to-reach population through social media and suggests new directions for future studies.

(JMIR Ment Health 2020;7(4):e17188) doi:10.2196/17188

KEYWORDS

loneliness; text postings; behavior online; social media; computer-based analysis; online self-disclosure

Introduction

Loneliness has become a public health problem described as an epidemic in modern society [1]. Recent research has produced evidence of the widespread nature of the condition and of the long-term impact of loneliness on health [2]. Loneliness poses risks to emotional and social health and to physical well-being. More than half (55%) of lonely people were more likely to be of poor health [3]. The prevalence of suicide ideation has been found to increase with the degree of loneliness [4,5].

Loneliness is defined as a subjective and emotional experience of unpleasant response to a lack of satisfactory companionship [6]. It appears across all age groups. More than a quarter of UK adults report sometimes being lonely, and 6% of them report being lonely all or most of the time [7]. In the 2002 Health and Retirement Survey, 19.3% of US adults older than 65 years

reported feeling lonely for much of the previous week [8]. A 2018 study of adults [9] found that those 18-22 years of age had the highest average loneliness score, 48.3%, of all the generations surveyed, followed by those 23-27 years of age with an average score of 45.3%. The lowest loneliness scores were among the oldest adults, those ≥72 years of age, with an average score of 38.6%.

Loneliness has been increasing in modern society [10], and there are concerns that using new online technologies and social media is contributing to rising loneliness. Previous investigations of the relationships between loneliness and media use suggested that lonely people do attempt to compensate by watching television, listening to radio, reading magazines, or going to movies [11]. However, it remains unclear whether social media provides the same relief from loneliness as these more traditional media.

Some studies suggest that loneliness can predict internet addiction [12] and problematic use of the internet [13], assuming that loneliness is an important determinant of social media use [14,15]. Dittmann [16] noted that loneliness appeared higher among participants who not only spent a lot of time online (more than 40 hours per week), but also preferred online interactions instead of face-to-face interactions or telephone communications. The heaviest internet users were more likely to demonstrate shyness, loneliness, and dissociation [17].

Loneliness has been found to be a positive predictor of Facebook addiction, standard Facebook use, and Facebook entertainment [18]. Passive social media use has also been found to be positively related to user's depressive symptoms [19]; social anxiety [20]; increased loneliness, envy, and shame [21]; and lower life satisfaction [22]. Lou et al [23] found that there was no significant difference in loneliness between those who create social media content and those who consume it—both creation and consumption were significantly related to loneliness. Błachnio et al [24] also found that people who had a compulsion to use their phone during contact with other people more often used Facebook in an excessive way and felt lonelier and scored lower on self-esteem and satisfaction with life.

Some researchers have cautioned against internet use since it creates loneliness, but others have identified its beneficial effects on social capital [25], well-being [26], and loneliness [27].

When confronted with excessive life stress, users may use social media as a means of stress relief or as a stress-coping strategy to escape from reality and to compensate for unsatisfactory social interactions [28]. A study of first-year students revealed that social media users seek out friends to dismiss the stress that is associated with poor adjustment [29]. Thus, social media could be used to satisfy users' need for psychological escape when confronted with real-life problems or challenging situations, especially to reduce emotional stress.

Pittman [30] found that individuals who used social media more were less likely to report being lonely. Gardner, Pickett, and Knowles [31] identified "social snacking behaviors" (such as looking at photos or rereading old emails) as symbolic social behaviors that can alleviate loneliness by serving as a reminder of existing social bonds. Clark et al [32] found that, although adolescents might be more socially (physically) isolated today, they also see less of a need for physical relationships than in the past. Previous generations grew up satisfying the need for social connection through physical relationships, but adolescents today appear to be more comfortable satisfying that need through digitally mediated activities [30].

As reviewed above, a large body of studies has been devoted to the increasing use of the internet, social media, and smartphones, and their effects on loneliness levels. Because of the diverse and even contradictory research results, such as those described above, there is a need to keep exploring the effects of digital behavior on loneliness, as these seem to have quite diverse relationships with measures of psychological well-being. Activity on social networks has been studied as a window to people's behavior in the physical world, especially related to health and medicine [33]. For example, De Choudhury et al [34] studied transitions from discussion of depression to those of suicidal ideation on Reddit, a popular social network. Furthermore, postings on social networks have been correlated with depression [35,36]. Other work investigated tracking people's experiences of adverse drug reactions using Twitter data [37]. Furthermore, Yom-Tov et al [38] demonstrated how physical and mental characteristics of people with anorexia could be tracked in online social networks through their postings and self-reported characteristics (both physical and mental). More generally, it has been suggested that digital traces can enrich and enhance psychological research [39] by providing data from people in their natural environment.

Discussions in online networks and online communities were also studied as a window to medical decision making. As such, several works [39,40] investigated an online forum for considerations of parents related to vaccinations. Similarly, Huber et al [41] examined discussion threads regarding treatment choices in a group for patients with prostate cancer.

The aim of this study is to expand the knowledge on the determinants and trends of loneliness via online postings in a social media forum. We provide an analysis of social media postings written by users who self-identify as lonely. Posting in a social media forum is understood here as an attempt to initiate social interaction online, and it is an integral and ongoing part of social media behavior. Thus, we first investigated whether loneliness can be assessed by an unrelated person that reads a post. Then, we studied the temporal trends of loneliness levels and their correlations with loneliness-related interest topics to understand the trajectory of posting behavior. We showed that loneliness, as expressed in social media postings, can be assessed by unrelated individuals. We developed an automated algorithm that can assess loneliness given the text of a posting and validated this algorithm both by comparing it to human-generated scores and by showing consistency in the scores of the same individuals. This enabled us to use the model to score a large number of postings on loneliness and to propose the following hypotheses.

- Hypothesis 1: Temporal trends in loneliness scores are associated with future topics of interest, including serious mental health outcomes and suicide.

- Hypothesis 2: A decrease in loneliness scores is associated with lower probability of suicide postings, compared to increasing or even constant scores.

Methods

Participants and Procedure

This study was an internet-based field experiment. To test our hypotheses, we chose data from the social network Reddit (reddit.com). Reddit is a social news aggregation, web content rating, and discussion website. Posts are organized by subject into user-created boards called "subreddits", which cover a variety of topics including news, science, movies, video games, music, books, fitness, image-sharing, and loneliness. As of March 2019, Reddit had 542 million monthly visitors (234 million unique users) and was ranked as the sixth most visited website in the United States and 21st in the world, according to Alexa Internet, with 53.9% of its user base coming from the United States, followed by the United Kingdom at 8.2%, and Canada at 6.3% [42].

Reddit has all the main features of discussion forums in that users post their own personal messages and photos. Like other major blog sites, anyone can visit other users' forums to read and post messages with simple clicks. Thus, everybody can post their personal messages knowing that others can easily access those messages. Each user is identified by a username, which can be (and often is) anonymous.

We extracted all postings made to the Reddit discussion board on loneliness (r/lonely) between January 1, 2015, and December 31, 2018 (4 years). We further extracted all posts on other subreddits made during these dates by people who posted on the loneliness discussion board. Posts made by people who subsequently deleted their account were removed from analysis.

On July 2014, a user posted a link to the long-form University of California, Los Angeles (UCLA) loneliness questionnaire and asked users to report their scores. A total of 40 users reported their scores, and, of those, 23 posted messages to r/lonely or to r/ForeverAlone (a similar subreddit). The long-form UCLA questionnaire is closely correlated with the UCLA Loneliness Scale (ULS-6) questionnaire used in our work (see Measures) [43]. We extracted these postings and reported the validity of our models by comparing self-reported scores to estimated scores, as described in the Measures sections.

Measures

We asked 7 crowdsourcing workers on Mechanical Turk to label a random sample of 236 posts (made by 230 users) from the loneliness discussion board. Workers were shown a post, asked to imagine how they would feel had they wrote this post, and then answer the short-form ULS-6 to measure respondents' loneliness [44]. The ULS-6 is a widely used tool [45] for gauging the loneliness of respondents. It includes six statements that people are asked to rate their agreement with. The total score is the aggregate of the agreement scores.

We then averaged the scores given by the workers for each question and calculated the loneliness score for each user. High scores indicate a higher degree of loneliness. Score reliability was estimated using interclass correlation and Krippendorff's alpha [46].

Using the 236 labeled posts, we built a predictive model for loneliness scores given the text of a question. All words, word pairs, and word triplets that appeared in 5 or more postings and fewer than 50% of postings were used as attributes for predicting loneliness scores. Additionally, we extracted sentiment attributes from the NLTK [47] package. This package provides an estimate of the positive, negative, neutral, and composite valence of sentiment.

The two feature families (phrases and sentiment) were modeled using a random forest with 50 trees. We noted that other models, including regression trees and linear models, achieved worst performance.

The random forest model was then applied to all postings from the loneliness forum, and further analysis (ie, stratifying posts according to predicted loneliness scores) was conducted.

All analysis was conducted using Matlab 2019a. This work was approved by the Institutional Review Board of the Technion – Israel Institute of Technology.

Results

A total of 19,668 postings from 11,054 users were found on the loneliness forum on Reddit. A further 406,714 postings were found to have been made by these users on other Reddit forums. Of the latter, 20.71% (84,249/406,714) were made after the first post in the loneliness forum and 79.23% (322,246/406,714) before it.

We first analyzed the reliability of unrelated individuals in estimating loneliness. Interclass correlation among crowdsourcing workers across the short-form ULS-6 was, on average, 0.19 (SD 0.03). Krippendorff's alpha was, on average, 0.19 (SD 0.03). This agreement is sufficient to allow a rough estimate of loneliness to be assessed by these workers. This lends support to our assumption.

The predictive model was trained separately for each of the ULS-6 statements using leave-one-out error estimation. The Pearson correlation between the total predicted score and the score estimated using crowdsourcing was 0.36 (P<.001). The use of phrases (word, word pairs, and triplets) alone reached a Pearson correlation of 0.21 (P=.001), and the use of sentiment attributes reached 0.34 (P<.001).

To estimate the validity of our models, we compared self-reported loneliness scores that were available for 23 users with the scores estimated by applying the models to their postings, as described in the Methods section. The score for users who posted more than once was taken as the maximum score across posts. The Spearman correlation between the estimated and reported loneliness scores was 0.48 (P=.02). Thus, our model represents moderate validity.

All collected posts in the loneliness forum were then scored using the model. Approximately 20.35% (2250/11,054) of users made more than one post in the loneliness forum. The correlation

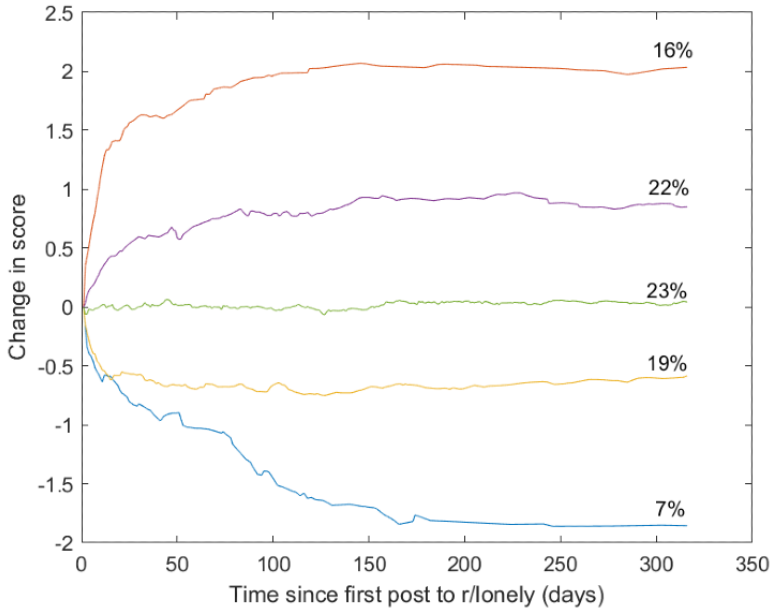
between the loneliness scores of the same user at different times is Spearman ρ=0.22 (*P*<.001).

Since both annotator agreement and classification accuracy were not high, we stratified users into five percentile bands, from people whose posts received an average predicted score in the low 20% of the scores, up to the top 20%.

We analyzed the trends in loneliness postings over time for 356 users who had five or more postings in the loneliness forum. First, their predicted loneliness scores were interpolated using linear interpolation from the first time they posted and for a period of 316 days, which represents the 50th percentile of the time span between the first and last posting among these users. These interpolated scores were normalized by removing the value of the first predicted score for each user and then clustered using k-means into 5 clusters, where each user was represented by their predicted loneliness score on each day.

Figure 1 shows the centroids of the 5 clusters as well as the percentage of users in each cluster who asked a question in the suicide watch forum (Multimedia Appendix 1 shows the figure with error bars). The average time between a posting to the loneliness forum and to the suicide watch forum was 215 days (SD 251). The partition is nonrandom (chi-square goodness of fit test, *P*<.001). As the figure shows, a decrease in loneliness scores is associated with a lower likelihood of suicide postings, compared to increasing or even constant scores. This lends support to our second hypothesis.

Figure 1. Centroids of users clustered over time by predicted loneliness scores and the percentages of users in each cluster who made posts in the suicide watch forum.



Notably, users who had a large increase in their loneliness were somewhat less likely to post on the suicide watch forum. However, in contrast to other clusters, all the users in this cluster posted in the suicide watch forum prior to their first post in the loneliness forum.

We noted that people with constant scores (the central cluster) also had high transition rates to r/SuicideWatch. We hypothesize that users with constant loneliness scores still suffer from emotional vulnerability that can affect their transition to suicidal ideation, see also [34].

We further noted that, in general, loneliness scores stabilized after around 100 days. We hypothesize that, after an initial period (reflected by beginning to post on r/lonely), people's loneliness scores stabilize and do not change significantly during the observation window.

To model the relationship between postings to the loneliness forum and future posting to the suicide watch forum, we examined users who made 2 or more posts to the loneliness forum and modeled future postings to the suicide watch forum using a Cox proportional-hazard model. The independent variables to this model were the current predicted loneliness score and the change in loneliness score. We included only cases where the time between postings on the loneliness forum was within 180 days.

Table 1 shows the model parameters for this analysis. We note that the time of day and day of the week were not statistically significantly associated with posting to the suicide forum. As the table shows, a lower current score is strongly associated with future postings to the suicide watch forum, as is a recent increase in loneliness scores. This provides support to our first hypothesis.

We modeled postings to the 50 most popular Reddit forums (excluding r/SuicideWatch analyzed above) among people who posted to the loneliness forum using the same methodology as above. Table 2 shows the model parameters for those forums where one or both variables were statistically significant at *P*<.05 using the Bonferroni correction.

Table 1. Cox proportional-hazard model for future postings on the suicide watch forum. N=6751.

Attribute	Exp (B)	SE	*P* value
Current score	0.775	0.102	.01
Change in score	1.195	0.079	.02

Table 2. Cox proportional-hazard model for future postings on different forums among the 50 most popular forums among people who posted in the loneliness foruma.

Forum	Exp (B)				Topics
	Current score	*P* value	Change in scores	*P* value	
r/depression	1.36	<.001	0.84	<.001	Depression
r/Drugs	1.44	<.001	0.80	.008	Illicit drugs

[a]Only forums where at least one variable was statistically significant (*P*<.05 with Bonferroni correction) are shown. The description column of topics is taken from the official description of the subreddit, or from the authors' understanding of this subreddit, if no official description exists.

We selected users who posted to the loneliness forum and to at least 10 or more other forums. We then analyzed posts made by these users to forums that had postings by at least 50 of these users and not more than 50% of them. Users were represented by the number of posts to each forum and clustered using nonnegative matrix factorization with 10 clusters.

Table 3 shows the most important forums in each cluster, according to the strength of the latent factor. Labels for the clusters are generalizations proposed by the authors.

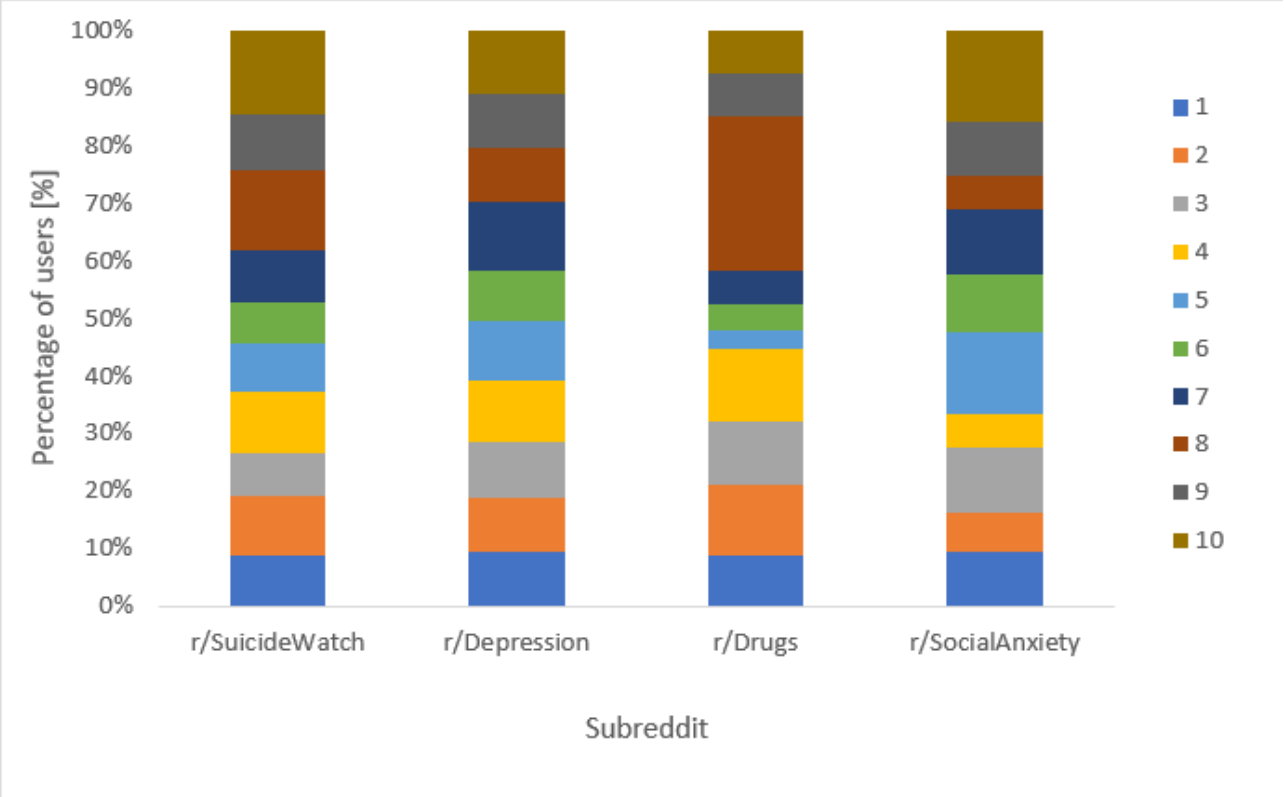
Interestingly, the likelihood of users in each of the clusters to post in subreddits related to the future behaviors shown in Table 2 varies significantly, as shown in Figure 2.

As the figure shows, people in cluster 8 were most likely to post about drugs. Together with people in cluster 10, they were also the most likely to post in the suicide watch subreddit. People in cluster 7 were the most likely to post in the depression subreddit. People in cluster 5 were low in all these behaviors. Thus, we deduce that people who post in the loneliness forum have diverse interests and backgrounds, and these behaviors are further associated to significantly varying degrees with interest in serious mental conditions.

Table 3. Most important forums in each cluster, according to the strength of the latent factors identified by nonnegative matrix factorization. Cluster titles are suggested by the authors.

Cluster (Title)	Subreddits
Cluster 1 (Depression)	• r/depression • r/depressed • r/misanthropy • r/disability • r/selfharm
Cluster 2 (Leisure)	• r/funny • r/weddingplanning • r/pics • r/Atlanta • r/gifs
Cluster 3 (Other)	• r/3amjokes • r/NoStupidQuestions • r/satanism • r/findareddit • r/Lightbulb
Cluster 4 (Gaming)	• r/leagueoflegends • r/summonerschool • r/TeamRedditTeams • r/LeagueConnect • r/formula1
Cluster 5 (Music)	• r/shittyideas • r/PipeTobacco • r/Showerthoughts • r/Blink182 • r/ToolBand
Cluster 6 (Young males)	• r/ForeverAlone • r/virgin • r/NEET • r/AnxietyDepression • r/Fucktheltright
Cluster 7 (Multimedia)	• r/GifSound • r/Supernatural • r/HFY • r/dvdcollection • r/FlashTV
Cluster 8 (Teens)	• r/teenagers • r/SonicTheHedgehog • r/nukedmemes • r/fakealbumcovers • r/nin
Cluster 9 (Gaming)	• r/Warthunder • r/hoi4 • r/pumparum • r/shittydarksouls • r/darksouls3
Cluster 10 (Illicit drugs)	• r/opiates • r/Stims • r/OpiatesRecovery • r/benzodiazepines • r/trees

Figure 2. Percentage of people in each of the clusters shown in Table 2 who posted in specific subreddits. The partitions are statistically significant (chi-square test) only for r/Drugs (*P*=.049).



Discussion

The aim of this study is to expand the knowledge on the determinants of loneliness via online postings in a social media forum. To the best of our knowledge, this is the first study to examine such relations on a large scale and with this methodology approach.

We found that there is a correlation between the total predicted loneliness score and the score estimated by crowdsourcing workers, as we assumed. The posted text can be read and assessed for its loneliness level to an acceptable degree by others, as if they were in the post writer's shoes. This is in lieu of a self-reported assessment by the post writers themselves, which is impossible to obtain in an anonymous social network. This is particularly important considering the subjective, dynamic, and ubiquitous features and bias of loneliness online and offline that could impact how people assess their own levels of loneliness.

Having demonstrated that loneliness-related postings can be assessed, we further demonstrated that an automated algorithm could assess loneliness given the text of a posting. We validated this algorithm both by comparing it to human-generated scores and by showing consistency in the scores of the same individuals. Additionally, for a small sample of users who provided an assessment of their scores in unrelated postings, good correlation was attained with predicted scores, demonstrating validity of the text-based estimator. This enabled us to use the model to score a large number of postings on loneliness and stratify users by their interests, analyzing a total of 19,668 postings from 11,054 users in the loneliness forum of Reddit.

Our results provide several interesting observations into the correlates of loneliness. First, loneliness in the context of online posting behavior is not a single-faceted phenomenon. Although our data was collected from a single subreddit, it has a trajectory that can be followed via posting behavior. Thus, people who posted on this forum could be clustered by their interests according to other subreddits where they were active, and these were correlated with specific outcomes. This clustering suggested different population characteristics, all who are lonely, and new variables to further explore the digital behavior of lonely people.

This study also confirmed that temporal trends in loneliness scores are associated with future topics of interest, including suicide. Our results show that changes in the degree of loneliness are strongly associated with suicidality (hazard ratio 1.19) and with other detrimental behaviors such as depression and illicit drug use. We confirmed that a decrease in loneliness scores is associated with a lower likelihood of suicide postings, compared to increasing or even constant scores.

We noted that the clustering seems to imply that a rise in loneliness scores is associated with a higher likelihood of posting activity on the suicide subreddit. However, whereas the Cox model measures short-term change, the clustering shows the long-term trend. Thus, we posit that long-term increase in loneliness as well as short-term reduction in loneliness are both associated with suicide ideation. We attribute the latter to the cathartic effect, whereby the ideation to suicide may cause a

reduction in negative emotions such as anxiety and, perhaps, loneliness.

This study has certain limitations that could be considered in future studies. First, we verified the validity of our measure of loneliness, obtained using crowdsourcing, through internal agreement between labelers. Future work will directly assess the ability of unrelated individuals to assess loneliness by comparing their scores with those of the individual who created the posting. Second, while crowdsourcing and the statistical model enabled an assessment of loneliness scores, future work will attempt to improve this assessment, either by using new methods of language processing or by including other information on users, such as demographics.

Another limitation of our study is in the use of a black box model based on word combinations. This makes it difficult to assess the reasons that a post is deemed to be associated with high or low levels of loneliness. Such an understanding could contribute to designing interventions, for example, responses that could reduce levels of loneliness and increase windows of opportunity between the first post on loneliness and the first post on suicide. Further research in this area should include gender, age, and other cultural variables, such as internet, social media and smartphone addiction, comparison between posting pictures and texts, and comparison between text posting in forums and other online activities such as shopping and information seeking. Another natural avenue for future work is an intervention (either manual or automatic) in the forum to reduce loneliness.

Despite these limitations, we believe that this research can improve our understanding of the impact of the internet and social media on social lives and the ubiquitous emotional phenomenon of loneliness. Our model can be compared across cultures.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Error bars.
[DOCX File , 284 KB - mental_v7i4e17188_app1.docx]

References

1. Killeen C. Loneliness: an epidemic in modern society. J Adv Nurs 1998 Oct;28(4):762-770. [doi: 10.1046/j.1365-2648.1998.00703.x]
2. Williams SE, Braun B. Loneliness and social isolation—a private problem, a public issue. J Fam Consum Sci 2019 Feb 01;111(1):7-14. [doi: 10.14307/JFCS111.1.7]
3. Anderson GO. AARP Research. 2010 Sep. Loneliness among older adults: a national survey of adults 45+ URL: https://www.aarp.org/research/topics/life/info-2014/loneliness_2010.html [accessed 2020-03-22]
4. Stravynski A, Boyer R. Loneliness in relation to suicide ideation and parasuicide: a population-wide study. Suicide Life Threat Behav 2001;31(1):32-40. [doi: 10.1521/suli.31.1.32.21312] [Medline: 11326767]
5. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
6. Heinrich LM, Gullone E. The clinical significance of loneliness: a literature review. Clin Psychol Rev 2006 Oct;26(6):695-718. [doi: 10.1016/j.cpr.2006.04.002] [Medline: 16952717]
7. Victor CR, Yang K. The prevalence of loneliness among adults: a case study of the United Kingdom. J Psychol 2012;146(1-2):85-104. [doi: 10.1080/00223980.2011.613875] [Medline: 22303614]
8. Theeke LA. Predictors of loneliness in U.S. adults over age sixty-five. Arch Psychiatr Nurs 2009 Oct;23(5):387-396. [doi: 10.1016/j.apnu.2008.11.002] [Medline: 19766930]
9. CIGNA. CIGNA. 2018. CIGNA Survey of 20,000 Americans examining behaviors driving loneliness in the United States URL: https://www.cigna.com/about-us/newsroom/studies-and-reports/loneliness-epidemic-america [accessed 2020-03-22]
10. Nowland R, Necka EA, Cacioppo JT. Loneliness and social internet use: pathways to reconnection in a digital world? Perspect Psychol Sci 2018 Jan;13(1):70-87. [doi: 10.1177/1745691617713052] [Medline: 28937910]
11. Rokach A. Surviving and coping with loneliness. J Psychol 1990 Jan;124(1):39-54. [doi: 10.1080/00223980.1990.10543204] [Medline: 2319485]
12. Özdemir Y, Kuzucu Y, Ak S. Depression, loneliness and Internet addiction: how important is low self-control? Computers in Human Behavior 2014 May;34:284-290. [doi: 10.1016/j.chb.2014.02.009]
13. Bian M, Leung L. Linking loneliness, shyness, smartphone addiction symptoms, and patterns of smartphone use to social capital. Social Science Computer Review 2014 Apr 08;33(1):61-79. [doi: 10.1177/0894439314528779]
14. Ye Y, Lin L. Examining relations between locus of control, loneliness, subjective well-being, and preference for online social interaction. Psychol Rep 2015 Feb;116(1):164-175. [doi: 10.2466/07.09.pr0.116k14w3]
15. Song H, Zmyslinski-Seelig A, Kim J, Drent A, Victor A, Omori K, et al. Does Facebook make you lonely?: a meta analysis. Computers in Human Behavior 2014 Jul;36:446-452. [doi: 10.1016/j.chb.2014.04.011]

16. Dittmann KL. A study of the relationship between loneliness and Internet use among university students. Dissertations 2003 [FREE Full text]
17. Kubey RW, Lavin MJ, Barrows JR. Internet use and collegiate academic performance decrements: early findings. Journal of Communication 2006. [doi: 10.1111/j.1460-2466.2001.tb02885.x]
18. Błachnio A, Przepiorka A, Boruch W, Bałakier E. Self-presentation styles, privacy, and loneliness as predictors of Facebook use in young people. Personality and Individual Differences 2016 May;94:26-31. [doi: 10.1016/j.paid.2015.12.051]
19. Lup K, Trub L, Rosenthal L. Instagram #instasad?: exploring associations among instagram use, depressive symptoms, negative social comparison, and strangers followed. Cyberpsychol Behav Soc Netw 2015 May;18(5):247-252. [doi: 10.1089/cyber.2014.0560] [Medline: 25965859]
20. Orchard LJ, Fullwood C, Galbraith N, Morris N. Individual differences as predictors of social networking. J Comput-Mediat Comm 2014 Feb 17;19(3):388-402. [doi: 10.1111/jcc4.12068]
21. Krasnova H, Widjaja T, Buxmann P, Wenninger H, Benbasat I. Research note—Why following friends can hurt you: an exploratory investigation of the effects of envy on social networking sites among college-age users. Information Systems Research 2015 Sep;26(3):585-605. [doi: 10.1287/isre.2015.0588]
22. Frison E, Eggermont S. Toward an integrated and differential approach to the relationships between loneliness, different types of Facebook use, and adolescents' depressed mood. Communication Research 2015 Dec 03:009365021561750. [doi: 10.1177/0093650215617506]
23. Lou LL, Yan Z, Nickerson A, McMorris R. An examination of the reciprocal relationship of loneliness and Facebook use among first-year college students. Journal of Educational Computing Research 2012 May 10;46(1):105-117. [doi: 10.2190/ec.46.1.e]
24. Błachnio A, Przepiorka A. Be aware! If you start using Facebook problematically you will feel lonely: phubbing, loneliness, self-esteem, and Facebook intrusion. A cross-sectional study. Social Science Computer Review 2018 Feb 18;37(2):270-278. [doi: 10.1177/0894439318754490]
25. Steinfield C, Ellison NB, Lampe C. Social capital, self-esteem, and use of online social network sites: a longitudinal analysis. Journal of Applied Developmental Psychology 2008 Nov;29(6):434-445. [doi: 10.1016/j.appdev.2008.07.002]
26. Valkenburg PM, Peter J. Internet communication and its relation to well-being: identifying some underlying mechanisms. Media Psychology 2007 Mar;9(1):43-58. [doi: 10.1080/15213260709336802]
27. Fokkema T, Knipscheer K. Escape loneliness by going digital: a quantitative and qualitative evaluation of a Dutch experiment in using ECT to overcome loneliness among older adults. Aging Ment Health 2007 Sep;11(5):496-504. [doi: 10.1080/13607860701366129] [Medline: 17882587]
28. Hou X, Wang H, Guo C, Gaskin J, Rost DH, Wang J. Psychological resilience can help combat the effect of stress on problematic social networking site usage. Personality and Individual Differences 2017 Apr;109:61-66. [doi: 10.1016/j.paid.2016.12.048]
29. Kalpidou M, Costin D, Morris J. The relationship between Facebook and the well-being of undergraduate college students. Cyberpsychol Behav Soc Netw 2011 Apr;14(4):183-189. [doi: 10.1089/cyber.2010.0061] [Medline: 21192765]
30. Pittman M. Creating, consuming, and connecting: examining the relationship between social media engagement and loneliness. The Journal of Social Media in Society 2015:4.
31. Gardner WL, Pickett CL, Knowles M. Social snacking and shielding: using social symbols, selves, and surrogates in the service of belonging needs. In: The Social Outcast: Ostracism, Social Exclusion, Rejection, and Bullying. 2005 Presented at: Sydney Symposium of Social Psychology series; 2005; Sydney, Australia p. 227-241. [doi: 10.4324/9780203942888]
32. Clark DMT, Loxton NJ, Tobin SJ. Declining loneliness over time: evidence from american colleges and high schools. Pers Soc Psychol Bull 2015 Jan 24;41(1):78-89. [doi: 10.1177/0146167214557007] [Medline: 25422313]
33. Yom-Tov E. Crowdsourced Health: How What You Do on the Internet Will Improve Medicine. Boston, MA, USA: MIT Press; 2016.
34. De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M. Discovering shifts to suicidal ideation from mental health content in social media. In: Proc SIGCHI Conf Hum Factor Comput Syst. 2016 May Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; 2016; San Jose, CA, USA p. 2098-2110 URL: http://europepmc.org/abstract/MED/29082385 [doi: 10.1145/2858036.2858207]
35. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. 2013 Presented at: Seventh International AAAI Conference on Weblogs and Social Media; 2013; Boston, MA, USA.
36. De Choudhury M, Counts S, Horvitz E, Hoff A. Characterizing and predicting postpartum depression from shared Facebook data. USA: ACM; 2014 Presented at: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing; 2014; Baltimore, Maryland, USA p. 626-638. [doi: 10.1145/2531602.2531675]
37. O'Connor K, Pimpalkhute P, Nikfarjam A, Ginn R, Smith KL, Gonzalez G. Pharmacovigilance on Twitter? Mining tweets for adverse drug reactions. AMIA Annu Symp Proc 2014;2014:924-933 [FREE Full text] [Medline: 25954400]
38. Yom-Tov E, Brunstein-Klomek A, Hadas A, Tamir O, Fennig S. Differences in physical status, mental state and online behavior of people in pro-anorexia web communities. Eat Behav 2016 Aug;22:109-112. [doi: 10.1016/j.eatbeh.2016.05.001] [Medline: 27183245]

39. Skea ZC, Entwistle VA, Watt I, Russell E. 'Avoiding harm to others' considerations in relation to parental measles, mumps and rubella (MMR) vaccination discussions - an analysis of an online chat forum. Soc Sci Med 2008 Nov;67(9):1382-1390. [doi: 10.1016/j.socscimed.2008.07.006] [Medline: 18703263]
40. Harmsen IA, Mollema L, Ruiter RA, Paulussen TG, de Melker HE, Kok G. Why parents refuse childhood vaccination: a qualitative study using online focus groups. BMC Public Health 2013 Dec 16;13(1):1183 [FREE Full text] [doi: 10.1186/1471-2458-13-1183] [Medline: 24341406]
41. Huber J, Ihrig A, Peters T, Huber CG, Kessler A, Hadaschik B, et al. Decision-making in localized prostate cancer: lessons learned from an online support group. BJU Int 2011 May;107(10):1570-1575 [FREE Full text] [doi: 10.1111/j.1464-410X.2010.09859.x] [Medline: 21105988]
42. Wikipedia. List of most popular websites URL: https://en.wikipedia.org/wiki/List_of_most_popular_websites [accessed 2019-07-02]
43. Wongpakaran N, Wongpakaran T, Pinyopornpanish M, Simcharoen S, Suradom C, Varnado P, et al. Development and validation of a 6-item Revised UCLA Loneliness Scale (RULS-6) using Rasch analysis. Br J Health Psychol 2020 Jan 30. [doi: 10.1111/bjhp.12404] [Medline: 31999891]
44. Nazzal FI, Cruz O, Neto F. Psychometric analysis of the short-form UCLA Loneliness Scale (ULS-6) among Palestinian university students. Interpersona 2018 Feb 23;11(2):113-125. [doi: 10.5964/ijpr.v11i2.269]
45. Amichai-Hamburger Y, Ben-Artzi E. Loneliness and Internet use. Computers in Human Behavior 2003 Jan;19(1):71-80. [doi: 10.1016/S0747-5632(02)00014-6]
46. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Communication Methods and Measures 2007 Apr;1(1):77-89. [doi: 10.1080/19312450709336664]
47. Natural Language Toolkit. URL: https://www.nltk.org/

Abbreviations

UCLA: University of California, Los Angeles
ULS-6: UCLA Loneliness Scale

Edited by G Eysenbach; submitted 25.11.19; peer-reviewed by Y Mejova, W Xu, S Ge; comments to author 06.01.20; revised version received 18.01.20; accepted 02.03.20; published 20.04.20.

Please cite as:
Mazuz K, Yom-Tov E
Analyzing Trends of Loneliness Through Large-Scale Analysis of Social Media Postings: Observational Study
JMIR Ment Health 2020;7(4):e17188
URL: http://mental.jmir.org/2020/4/e17188/
doi:10.2196/17188
PMID:32310141

Original Paper

Portuguese Psychologists' Attitudes Toward Internet Interventions: Exploratory Cross-Sectional Study

Cristina Mendes-Santos[1,2], MSc; Elisabete Weiderpass[3], MSc, PhD; Rui Santana[2], PhD; Gerhard Andersson[4,5], PhD

[1]Department of Culture and Society, Linköping University, Linköping, Sweden

[2]NOVA National School of Public Health, Public Health Research Centre, Universidade Nova de Lisboa, Lisboa, Portugal

[3]International Agency for Research on Cancer, Lyon, France

[4]Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

[5]Department of Clinical Neuroscience, Psychiatry Section, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Cristina Mendes-Santos, MSc
Department of Culture and Society
Linköping University
Linköping
Sweden
Phone: 46 13 28 10 0
Email: cristina.mendes.santos@liu.se

Abstract

Background: Despite the significant body of evidence on the efficacy and cost-effectiveness of internet interventions, the implementation of such programs in Portugal is virtually non-existent. In addition, Portuguese psychologists' use and their attitudes towards such interventions is largely unknown.

Objective: The aim of this study was to explore Portuguese psychologists' knowledge, training, use and attitudes towards internet interventions; to investigate perceived advantages and limitations of such interventions; identify potential drivers and barriers impacting implementation; and study potential factors associated to previous use and attitudes towards internet interventions.

Methods: An online cross-sectional survey was developed by the authors and disseminated by the Portuguese Psychologists Association to its members.

Results: A total of 1077 members of the Portuguese Psychologists Association responded to the questionnaire between November 2018 and February 2019. Of these, 37.2% (N=363) were familiar with internet interventions and 19.2% (N=188) considered having the necessary training to work within the field. 29.6% (N=319) of participants reported to have used some form of digital technology to deliver care in the past. Telephone (23.8%; N=256), e-mail (16.2%; N=175) and SMS (16.1%; N=173) services were among the most adopted forms of digital technology, while guided (1.3%; N=14) and unguided (1.5%; N=16) internet interventions were rarely used. Accessibility (79.9%; N=860), convenience (45.7%; N=492) and cost-effectiveness (45.5%; N=490) were considered the most important advantages of internet interventions. Conversely, ethical concerns (40.7%; N=438), client's ICT illiteracy (43.2%; N=465) and negative attitudes towards internet interventions (37%; N=398) were identified as the main limitations. An assessment of participants attitudes towards internet interventions revealed a slightly negative/neutral stance (Median=46.21; SD=15.06) and revealed greater acceptability towards blended treatment interventions (62.9%; N=615) when compared to standalone internet interventions (18.6%; N=181). Significant associations were found between knowledge (χ^2_4=90.4; P<.001), training (χ^2_4=94.6; P<.001), attitudes (χ^2_3=38.4; P<.001) and previous use of internet interventions and between knowledge (χ^2_{12}=109.7; P<.001), training (χ^2_{12}=64.7; P<.001) and attitudes towards such interventions, with psychologists reporting to be ignorant and not having adequate training in the field, being more likely to present more negative attitudes towards these interventions and not having prior experience in its implementation.

Conclusions: This study revealed that most Portuguese psychologists are not familiar with and have no training or prior experience using internet interventions and had a slightly negative/neutral attitude towards such interventions. There was greater acceptability towards blended treatment interventions compared to standalone internet interventions. Lack of knowledge and training were identified as the main barriers to overcome, underlining the need of promoting awareness and training initiatives to ensure internet interventions successful implementation.

(JMIR Ment Health 2020;7(4):e16817) doi:10.2196/16817

KEYWORDS

attitudes; psychologists; e-mental health; internet interventions; Attitudes Toward Internet Interventions Survey (ATIIS); Portugal; EU

Introduction

Background

Advances in digital technology are transforming the current health care delivery paradigm, enabling health systems to overcome physical and organizational barriers and creating an opportunity to deliver accessible and convenient mental health care services at a distance [1]. In recent years, the development of internet interventions—self-help guided or unguided interventions based on established psychotherapy models operated via secure platforms or mobile apps that aim at providing synchronous or asynchronous health and mental health–related assistance—has generated significant evidence of efficacy and cost effectiveness [2-9]. There are several advantages associated with implementing internet interventions [10-15]: (1) low-threshold accessibility and dissemination potential, (2) high use flexibility and adaptability, (3) standardized structure and integrated treatment monitoring, (4) self-efficacy and patient empowerment promotion, (5) high level of anonymity and privacy, and (6) low delivery costs.

Despite evidence of efficacy and potential usefulness, implementation of internet interventions in clinical practice has been peculiarly slow [16]. In Portugal, a country characterized by a significant mental health treatment gap where treatment median delay reaches 23 years for anxiety disorders [17], internet interventions could represent an effective opportunity to reach those in need. Yet such interventions are virtually nonexistent. This fact may be related to potential barriers and limitations of internet interventions such as the absence of an adequate legal and regulatory framework for providing mental health care via the internet, health care professional and patient information and communication technologies (ICT) illiteracy, possible technological problems, potential security breaches associated with computerized systems, and health care provider and patient attitudes toward internet interventions [13]. As potential end users and prescribers of a variety of e-mental health programs (eg, guided and unguided internet interventions, apps, serious games), psychologists might play a crucial role in the uptake and dissemination of internet interventions. Thus, understanding attitudes toward such programs is key to identify drivers and barriers to implementation and overcome limitations to dissemination.

Prior Work

Previous studies have been performed to investigate psychologist attitudes toward internet interventions and categorize drivers and barriers to the adoption of such programs. Overall, findings suggest that therapist attitudes range from neutral and cautiously positive to generally positive [18], and several factors impact adoption.

In a study performed by Mora and colleagues [19] querying members of the New York State Psychological Association, theoretical orientation was found to predict the use of internet-based interventions. Dynamic and existential oriented therapists were less likely to endorse internet interventions and more likely to have negative attitudes toward them than therapists from other theoretical stances such as cognitive behavioral, cognitive, behavioral, and systems therapists, a finding that has been corroborated in other studies [20-24]. Another study performed by Simms and colleagues [25], focusing on Canadian mental health professional perceptions of telemental health, found overall positive attitudes particularly for clients in remote and rural locations and people with disabilities, and factors predicting frequency of use of telemental health were having performed previous training in the field, working within mental health for longer, and considering technology as easy to use. Corroborating these findings, a study by Bruno et al [26] focusing on Australian health professional attitudes reported that professionals with higher perceptions of usefulness and ease of use of internet-supported psychological interventions presented more positive attitudes toward using such interventions than professionals with lower perceptions of usefulness and ease of use. Moreover, the possibility of encouraging clients to develop self-management skills and reaching clients who might not otherwise engage in therapy were considered the main benefits of internet-supported psychological interventions by participants in this study. More recently, Feijt et al [15], in a qualitative study involving Dutch psychologists, found them to believe e-mental health brings new treatment possibilities (eg, virtual reality and biofeedback) and may accelerate the treatment process (mediated contact in-between regular sessions may intensify treatment and allows for the introduction of new therapeutic elements earlier in the process), reinforcing intimacy in the therapeutic relationship.

Contrasting with these findings, a more guarded attitude was identified in other studies, and major concerns related to security and ethical and legal requirements have been reported as important barriers to adoption [15,21,27-29]. Lack of clarity and knowledge regarding ethical and regulatory requirements emerged as important limitations in studies by Perle et al [21] and Glueckauf et al [27], and a possible threat to confidentiality and therapeutic boundaries posed by online communication was identified by Evans et al [28]. A further concern expressed in this study was that the therapeutic alliance could be negatively affected due to missing nonverbal cues in communication. This was also a finding in other studies [19,29].

Technical barriers such as connection challenges and disruptions were also flagged as potential limitations of internet interventions, and a major concern reported in several studies [19,21,27] relates to handling emergencies and managing crisis situations (eg, suicide ideation, child abuse) in the context of online practice. In a study by Glueckauf et al [27], over half of respondents reported inadequate skills in managing crisis situations in this environment. In addition, Perle and colleagues

[21] reported that adoption of telehealth depended on disorder type and was rejected for disorders considered difficult to treat even face-to-face, such as schizophrenia. Likewise, Vigerland et al [22] found Swedish mental health professionals to be strong supporters of computerized cognitive behavior therapy for preventing and treating mild to moderate problems, but more caution was reported regarding severe mental health problems.

Other important findings reported in previous research as influencing therapist attitudes relate to demographic and background factors (eg, age, gender, length of professional career, personal experience in using modern technologies) [19,21,30], practical concerns (eg, costs of setting up and maintaining the necessary infrastructure) [15], fear of replacement [31,32], social influence [33], skepticism about feasibility of delivery within existing care services [32,34], forces within the care system, design and usability [15], low patient engagement, and difficulties in managing comorbidities [32].

In Portugal, only one study addressed psychologist attitudes toward electronic psychological interventions (EPI). In the study, Neves et al [24] showed that Portuguese psychologists reported moderately less favorable attitudes toward EPI. Moreover, years of termination of vocational training and cognitive behavioral and eclectic/integrative theoretical orientations were predictors of positive attitudes toward EPI. Nevertheless, a full report on the findings of this study is unavailable for consultation, and gaps persist in our knowledge regarding Portuguese psychologists' use and attitudes toward internet interventions, perceived advantages and limitations of such interventions, drivers and barriers impacting implementation, and factors associated to adoption of internet interventions in the country.

Aim

In spite of the findings that psychologists' attitudes toward internet interventions appear to be positive, there are variations between countries [34] and lack of consensus between studies regarding factors influencing adoption. The aim of this study was to explore Portuguese psychologists' knowledge, training, use and attitudes toward internet interventions; to investigate perceived advantages and limitations of such interventions; identify potential drivers and barriers impacting implementation; and study potential factors associated to previous use and attitudes toward internet interventions.

Methods

Study Design and Procedures

This study was conducted in the framework of the iNNOVBC (A Guided Internet-Delivered Individually Tailored ACT-Influenced Cognitive Behavioral Intervention to Improve Psychosocial Outcomes in Breast Cancer Survivors) project (ClinicalTrials.gov NCT03275727) [35] and approved by the Portuguese Data Protection Committee (approval number: 10727/2017) and Portuguese Psychologists Association (Ordem dos Psicólogos Portugueses, or OPP) ethical committee, adopting an exploratory cross-sectional design. An anonymous online self-report questionnaire located on the Web-based survey platform LimeSurvey [36] was disseminated by OPP via email to its members. As being registered with OPP is a requirement to practice psychology in Portugal and email is the institutional channel of communication with its members, the whole universe of licensed psychologists was reached using this method. Participants initially accessed an introduction to the study and informed consent form, followed by a link providing access to the questionnaire. No follow-up reminders were sent to recipients of the questionnaire. The time frame of data collection was November 2018 to February 2019. After conclusion of the recruitment period, researchers exported data from LimeSurvey [36] to SPSS Statistics version 24.0 (IBM Corp).

Survey Development and Design

Due to the scarcity of adequate instruments designed to evaluate the outlined issues in the target population, the Attitudes Toward Internet Interventions Survey (ATIIS) was developed. After a comprehensive literature review was performed, most relevant publications were identified and served as a basis for development [20,21,26,34,37-40]. A preliminary version of this self-report questionnaire was created, pilot-tested with 3 participants, and subsequently checked by researchers and clinicians with some experience within the field of internet interventions and OPP. Changes were made in line with suggestions emerging from this process, and final item selection was completed by the authors. Selection criteria were redundancy, relevance of items, and face validity.

The final version of the survey comprised 38 items assessing 4 main categories: (1) information relating to frequency of use of digital technology and internet interventions in practice (eg, use and frequency of use of digital technology, provision and prescription of internet interventions, contexts and purposes of use); (2) knowledge and training within the internet interventions field; (3) perceived advantages and limitations of internet interventions and potential barriers and challenges impacting implementation; and (4) attitudes toward internet interventions (eg, related to efficacy and efficiency; privacy, security and confidentiality; patient empowerment and increased disinhibition; therapeutic processes and alliance; and blended, complementary, and stand-alone interventions). The attitudes section was composed of 21 items aimed at capturing cognitive, affective, and behavioral predispositions of favor or disfavor [41] toward internet interventions.

Demographic and background items were added to the questionnaire to gather supplementary information (eg, age, gender, educational and professional background, professional experience, and theoretical orientation). The survey questions were asked in the form of dichotomous and multiple choice questions and in the form of 5-point (0=completely disagree to 5=completely agree) Likert scales. Since it was not expected for participants to be familiar with the concept of internet interventions, an explanation of the concept based on the definition by Barak et al [42] was provided in the instructions section of the questionnaire.

An assessment of the validity of the attitudes section of the questionnaire resulted in 21 items clustering in two dimensions labeled as positive attitudes and negative attitudes. Reliability of the scale was also tested and considered excellent (α=.91).

A detailed description of the psychometric properties assessment process of this scale and its results can be found below. A copy of the instrument is available in Multimedia Appendix 1.

Statistical Analysis

Statistical analyses were divided into 4 steps and conducted using SPSS Statistics. First, descriptive statistics such as frequency distributions, measures of variability, and measures of central tendency were calculated to characterize the study sample and determine its face validity. These statistics encompassed demographic and background characteristics such as, age, gender, educational and professional background, professional experience (in years), and theoretical orientation.

Second, a psychometric properties evaluation process of ATIIS took place, and an exploratory factor analysis (EFA) based on the principal component analysis method using a varimax rotation was conducted to determine the factor structure of the questionnaire, perform scale purification, and determine the questionnaire's construct validity. The whole study sample (1077) was used for this purpose. The Kaiser-Meyer-Olkin (KMO) test and a Bartlett test of sphericity were calculated to measure sampling adequacy (confirmed if KMO value greater than .5) and appropriateness of the extracted factors (significant at $P<.05$), respectively. The initial model hypothesized that items would load on either a positive or a negative factor, and items with factor loadings above .40 were considered acceptable [43]. Scores on the negative items were reversed, and dimension scores were weighted, summed, and rescaled on a 100-point scale to simplify interpretation and obtain a continuous indicator of attitude toward internet interventions. Higher scores indicated a more positive attitude. The final version of ATIIS was then subject to a reliability analysis based on the computation of internal consistency (Cronbach alpha).

Following this process, results pertaining to frequency of use of digital technology and internet interventions in daily practice, provision and prescription of internet interventions, contexts and purposes of use, perceived advantages and limitations of internet interventions, and potential barriers and challenges impacting implementation were analyzed. Although some of the questionnaire items contained multiple response options for which up to 3 response categories could be selected, only single response options (eg, percentage of psychologists using chat services) rather than combined response options (eg, percentage of psychologists using chat services and videoconference) were calculated in order to simplify the analysis.

Finally, psychologist attitudes toward internet interventions were examined using descriptive statistics, and chi-square analysis and post hoc tests were used to determine if demographic (eg, sex and age) and background factors (eg, academic background, work context, years of professional experience, theoretical orientation), knowledge, training, previous experience of use, recommendation, future use, and attitudes toward internet interventions would be associated and differed between participants holding extreme attitudes toward internet interventions.

Results

Participants and Recruitment

The total sample comprised 1077 members of the OPP recruited between November 2018 and February 2019. Considering the number of psychologists registered as members at the time (21,214, data provided by T Pereira, OPP's head of cabinet), response rate was 5.08%. Although we cannot determine the representativeness of the sample (OPP's members demographic and background information is not available for consultation), demographic characteristics are similar to those published in the last census performed by OPP [44]. In this census, the mean age of Portuguese psychologists was 38 years, and 84.2% were female. On average, psychologists had 11 years of professional experience, and the majority held a license and/or master's degree. Only 7% held a doctoral degree.

In our study sample, 91.6% (987/1077) of respondents were female, and age ranged from 20 to 77 years (mean 38.21; SD 9.49 years). Most participants held a license and/or master's degree (722/1077, 67.0%), followed by postgraduate (273/1077, 25.3%), doctoral (75/1077, 7%), and bachelor's degrees (7/1077, 0.6%). The majority of participants were active (986/1077, 91.6%) and worked primarily in private practice (270/1077, 25.1%), educational/research institutions (252/1077, 23.4%), and charities/nonprofit organizations (208/1077, 19.3%). Only 6.3% (68/1077), 4.4% (47/1077), and 1.6% (17/1077) of psychologists worked in the National Health Service (NHS) at primary, secondary and tertiary care, respectively. As for the length of time working within the field of psychology, the sample was evenly distributed, with 12.6% (136/1077) of professionals working for less than a year; 21.0% (226/1077) practicing psychology from 2 to 5 years, 17.3% (186/1077) working between 6 to 10 years in the field, 20.0% (215/1077) practicing between 11 to 15 years, 15.4% (166/1077) working from 16 to 20 years in this domain, and 13.7% (148/1077) practicing psychology for more than 21 years. Cognitive behavioral therapy was the most common theoretical orientation (56.0%, 603/1077), with psychodynamic (14.8%, 159/1077) and eclectic (13.5%, 145/1077) orientations being second and third.

Attitudes Toward Internet Interventions Survey Psychometric Properties Assessment

In order to test the psychometric properties of the attitudes section of ATIIS, we explored its construct validity and reliability.

Construct Validity

An EFA based on principal component analysis and using a varimax rotation was conducted with the purpose of finding the underlying latent factors of ATIIS and determining the questionnaire's construct validity. The whole study sample (N=1077) was used in this analysis. A KMO=.93 confirmed the sampling adequacy, and a Bartlett test of sphericity, $\chi^2_{210}=8003.39$ ($P<.001$), indicated a possible statistically significant interrelationship between variables and, therefore, confirmed the factorial analysis validity to perform factor reduction.

The initial EFA resulted in 4 factors with eigenvalues above a Kaiser criterion of 1. However, a scree plot analysis revealed inflexions compatible with the retention of two factors. Due to convergence with theory, two factors were retained for the final EFA. The initial model hypothesized that items would load on either a positive or a negative factor. Total variance explained by these two factors was 44.10% (unrotated solution: factor one 36.2% and factor two 7.97% or rotated solution: factor one 22.40% and factor two 21.73%), and items clustering on these two factors suggested that the questionnaire measures two dimensions, labeled as positive attitudes (range of factor loadings: .375-.712) and negative attitudes (range of factor loadings: .459-.708). Items with factor loadings above r=.4 were considered as acceptable [43]. Scores on the negative items were reversed, and dimension scores were weighted, summed, and rescaled on a 100-point scale to simplify interpretation and obtain a continuous indicator of attitude toward internet interventions. Higher scores indicated a more positive attitude. A copy of the final questionnaire and item loading factors is presented in Multimedia Appendix 1.

Reliability

ATIIS reliability was assessed via the computation of Cronbach alpha. ATIIS total scale revealed excellent (α=.91) internal consistency and its subscales, positive (.88) and negative (.82) attitudes, showed good internal consistency [45].

Portuguese Psychologists Reported Knowledge About Internet Interventions

An examination of collected data indicated that 37.2% (363/978) of respondents were familiar with the concept of providing psychological support via the internet. Nevertheless, a narrower group reported knowing how these types of interventions work (218/978, 22.3%), and only 19.2% (188/978) were considered to have the necessary training to work in the field.

Frequency of Use of Digital Technology and Internet Interventions in Daily Practice

Around 29.6% (319/1077) of participants reported that they use or have used in the past some form of digital technology to provide support in the context of their practice. Of nonusers (758/1077, 70.4%), 61.7% (468/758) reported to be considering using it in the future. Telephone (256/1077, 23.8%), email (175/1077, 16.2%,) and short message service (SMS) or text message (173/1077, 16.1%) services were among the most used forms of digital technology, while chat services (66/1077, 6.1%) and unguided (16/1077, 1.5%) and guided (14/1077, 1.3%) internet interventions were much less used. However, 8.7% (94/1077) reported using videoconference services. Digital technology was mostly used by clinical and health psychologists (269/319, 84.3%), followed by educational psychologists (31/319, 9.7%). In most cases, digital technology was used as a complement to face-to-face interventions (288/319, 90.3%) rather than as a stand-alone interventions (31/319, 9.7%) for the purpose of treating mental health disorders such as anxiety or depression (205/319, 64.3%). Increasing accessibility to information and psychological care was reported as the main reason for using digital technology in practice by 54.5% (174/319) of respondents whereas only 0.6% (2/319) used it for research (see Table 1).

Table 1. Motivations for previous use of digital technology in psychological practice (n=319).

Motivation	Value, n (%)
Increasing accessibility to information and psychological care	174 (54.5)
Lowering the costs of psychological interventions	3 (0.9)
Increasing adherence to psychological interventions	61 (19.1)
Monitoring treatment progress	41 (12.9)
Facilitating follow-up care	26 (8.2)
Managing crisis situations	5 (1.6)
Improving career prospects	5 (1.6)
Research	2 (0.6)
Other	2 (0.6)

Almost a fifth (19.1%, 206/1077) reported that they recommend or have recommended in the past to their clients accessing online services or resources with the aim of improving their emotional wellbeing and/or health status. Most frequently recommended resources were websites providing information about mental and/or somatic health (57.3%, 118/206), blogs, discussion forums and social networks (38.3%, 79/206), videoconference-delivered psychological interventions (29.1%, 60/206), and apps (28.2%, 58/206). On the other hand, online support groups (21.8, 45/206) and guided (13.1%, 27/206) and unguided (4.4%, 9/206) internet interventions were the least recommended. Only a minority of respondents (3.7%, 39/1077) provided or recommended internet interventions to their clients in a regular basis.

Advantages and Limitations Associated With Internet Interventions

Considering the potential advantages of internet interventions, accessibility (860/1077, 79.9%), convenience (492/1077, 45.7%), and cost effectiveness (490/1077, 45.5%) of such interventions were considered the most important advantages. Conversely, ethical concerns (438/1077, 40.7%), client ICT illiteracy (465/1077, 43.2%), and client negative attitudes toward internet interventions (398/1077, 37.0%) were identified as the

main limitations. Other advantages and limitations associated with internet interventions are presented in Table 2.

Table 2. Advantages and limitations associated with internet interventions (n=1077).

Characteristic	Value, n (%)
Advantage	
Accessibility	860 (79.9)
Convenience	492 (45.7)
Economical (cost effectiveness and sustainability to health care systems)	490 (45.5)
Reduced stigma associated with psychological support/confidentiality	206 (19.1)
Privacy/anonymity	164 (15.2)
Health equity	142 (13.2)
Client empowerment	131 (12.2)
Personalized health care	89 (8.3)
None	77 (7.1)
Scientific evidence	43 (4)
Limitations	
Client information and communications technologies illiteracy	465 (43.2)
Ethical	438 (40.7)
Client attitudes toward internet interventions	398 (37.0)
Information systems security	386 (35.8)
Cultural	271 (25.5)
Therapist attitudes toward internet interventions	259 (24.0)
Health care systems not ready for implementation	234 (21.7)
Cost and accessibility to digital technology	144 (13.4)
Therapist information and communications technologies illiteracy	96 (8.9)
Other	82 (7.5)
Political (decision makers not interested in implementation)	50 (4.6)
Economical (cost effectiveness and sustainability to health care systems)	18 (1.7)
None	18 (1.7)

When questioned about the possibility of internet interventions presenting more disadvantages than advantages, only 24.5% (239/1077) of participants refuted this claim.

Barriers to Implementation of Internet Interventions

The main barriers to overcome in the implementation of internet interventions were related to limitations on the conceptual comprehension and implementation of self-help techniques by clients (676/1077, 62.8%), therapist perceptions of insufficient scientific evidence on the efficacy and cost effectiveness of internet interventions (670/1077, 62.2%), limitations on the adaptation of treatment protocols (665/1077, 61.7%), patient ICT illiteracy (516/1077, 47.9%), and low adherence both from patients (466/1077, 43.3%) and psychologists (437/1077, 40.6%) toward such programs. Negative attitudes presented both by clients (417/1077, 38.7%) and therapists (416/1077, 38.6%) were also considered an important obstacle to overcome in the implementation of internet interventions (see Table 3).

Table 3. Barriers to implementation of internet interventions (n=1077).

Characteristic	Value, n (%)
Faced by clients	
Ability to comprehend concepts and learn self-help techniques	676 (62.8)
Client information and communications technologies illiteracy	516 (47.9)
Low adherence	466 (43.3)
Negative attitudes	417 (38.7)
Scientific evidence (efficacy and cost effectiveness)	362 (33.6)
Costs and access to digital technology and information technology infrastructures	211 (19.6)
Time consumption	28 (2.6)
None	21 (1.9)
Faced by therapists	
Scientific evidence (efficacy and cost effectiveness)	670 (66.2)
Adaptation of treatment protocols to the digital environment	665 (61.7)
Low adherence	437 (40.6)
Negative attitudes	416 (38.6)
Clinician information and communications technologies illiteracy	178 (16.5)
Costs and access to digital technology and information technology infrastructures	91 (8.4)
Time consumption	38 (3.5)
None	20 (1.9)

Analysis of Portuguese Psychologist Attitudes Toward Internet Interventions

The median score on the ATIIS scale was 46.21 (SD 15.06), which corresponds to a slightly negative/neutral attitude toward internet interventions. Factors contributing to this predisposition relate to possible security (417/1077, 42.6%) and confidentiality (494/1077, 50.5%) breaches when using internet interventions, reported discomfort about dealing with sensitive information online (466/1077, 47.7%), perceived inaccuracy of remote psychological assessment processes (578/1077, 59.1%), perceived unsuitability of internet interventions for crisis management (473/1077, 48.4%), a disbelief on the possibility of establishing therapeutic alliance via the internet (356/1077, 36.5%), and a generalized perception of face-to-face interventions as being superior for client education/self-management skills development (702/1077, 71.8%) and mental disorders treatment (699/1077, 71.6%) compared with internet interventions. Absence of knowledge about the efficacy (561/1077, 57.4%) and efficiency (443/1077, 45.3%) of internet interventions and its impact on patient empowerment (461/1077, 47.2%) and a possible loss of control of the therapeutic process by clinicians (372/1077, 38.1%) also seem to influence this stance. Nevertheless, perceived convenience (473/1077, 45.6%) of internet interventions, encouragement of emotional expression in some cases (360/1077, 36.8%), facilitation of the follow-up process (440/1077, 45%), and the possibility of delivering blended (615/978, 62.9%) and pharmacotherapy complementary interventions (511/978, 52.2%) rather than stand-alone internet interventions (181/978, 18.6%) seem to balance attitudes regarding this matter.

Factors Associated With Previous Experience of Use and Attitudes Toward Internet Interventions

Chi-square tests and post hoc analyses were performed to examine possible associations between demographic factors (sex, age), background factors (academic background, work context, years of professional experience, theoretical orientation), knowledge, training, recommendation, future use, previous experience of use, and attitudes toward internet interventions. Differences in responses between participants with or without prior experience of use and holding extreme attitudes toward these interventions were also evaluated. Percentiles (assumed here as the percentage of scores that fall below the scores of interest) were computed to categorize participant attitudes and identify those who held extremely negative (scores <Q1=36.03, 244; 25.1%) and extremely positive (scores >Q3=53.49, 242; 24.9%) attitudes toward internet interventions.

Chi-square analyses (see Table 4) revealed a significant association between previous experience of use and age, theoretical orientation, work context, years of professional experience, knowledge, training, recommendation, and attitudes toward internet interventions.

Considering age, psychologists aged between 41 and 60 years were more likely to have used the telephone or internet to provide psychological support in the past, while psychologists aged 30 years and younger were less likely to have done it. Similarly, psychologists with less than 5 years of professional

experience were less likely to have already used such interventions, whereas psychologists with more than 16 years of professional experience were more likely to have used internet interventions in the past. Work context also seems to impact the use of internet interventions. Participants working at the NHS and in private practice had a higher probability of using the internet and telephone to provide care. Working at public services, education/research facilities, and charities made it less probable participants had adopted these interventions.

Regarding self-reported knowledge and training on internet interventions, psychologists reporting moderate to high knowledge and training were more likely to have prior experience in implementing such programs than those whom reported little to no knowledge about internet interventions. Furthermore, having a psychodynamic theoretical orientation impacted use positively, making it more likely that psychodynamic psychologists had used internet interventions in the past than expected. No significant associations were found between other theoretical stances and internet intervention adoption.

Extreme attitudes toward internet interventions seem, as well, to have significantly impacted adoption. Psychologists presenting more negative attitudes toward these interventions were less likely to have prior experience using internet interventions than expected and when compared with psychologists holding more positive attitudes. Finally, prior experience implementing internet interventions significantly affected referrals and the possibility of psychologists recommending such programs to their clients. Psychologists with prior experience of use were more likely to recommend internet interventions and online resources with the purpose of improving their clients' health status.

Table 4. Factors associated to previous experience of use.

Characteristic	Previous experience of use[a]		Chi-square tests		
	No	Yes	*P* value	Chi-square	Cramér V Φc
Age in years (n=1077)			**<.001**	**χ^2_4=42.4**	**.20**
≤30	207 (4.4)	47 (–4.4)	—	—	—
31-40	323 (1.4)	121 (–1.4)	—	—	—
41-50	152 (–3.5)	95 (3.5)	—	—	—
51-60	59 (–4.2)	52 (4.2)	—	—	—
≥61	17 (1.1)	4 (–1.1)	—	—	—
Theoretical orientation (n=1077)			**<.001**	**χ^2_6=27.7**	**.16**
Cognitive behavior therapy	435 (1.4)	168 (–1.4)	—	—	—
Psychodynamic	99 (–2.4)	60 (2.4)	—	—	—
Humanist	25 (–1.6)	17 (1.6)	—	—	—
Eclectic	93 (–1.8)	52 (1.8)	—	—	—
Systemic	24 (1.5)	5 (–1.5)	—	—	—
Other	37 (0.1)	15 (–0.1)	—	—	—
None	45 (3.9)	2 (–3.9)	—	—	—
Work context (n=1077)			**<.001**	**χ^2_7=50.9**	**.22**
National Health Service	82 (–2.2)	50 (2.2)	—	—	—
Private practice	151 (–6.0)	119 (6.0)	—	—	—
Public services	59 (2.4)	12 (–2.4)	—	—	—
Private companies	22 (1.0)	6 (–1.0)	—	—	—
Rehabilitation services/prisons	21 (0.9)	6 (–0.9)	—	—	—
Education/research institutions	190 (2.0)	62 (–2.0)	—	—	—
Charities	164 (3.0)	44 (–3.0)	—	—	—
Other	69 (1.5)	20 (–1.5)	—	—	—
Professional experience in years (n=1077)			**<.001**	**χ^2_7=49.0**	**.21**
≤1	122 (5.3)	14 (–5.3)	—	—	—
2-5	172 (2.1)	54 (–2.1)	—	—	—
6-10	129 (–0.3)	57 (0.3)	—	—	—
11-15	151 (–0.1)	64 (0.1)	—	—	—
16-20	96 (–3.9)	70 (3.9)	—	—	—
≥21	88 (–3.1)	60 (3.1)	—	—	—
Self-reported knowledge[a] (n=978)			**<.001**	**χ^2_4=90.4**	**.30**
Completely disagree	145 (6.1)	15 (–6.1)	—	—	—
Moderately disagree	200 (3.6)	52 (–3.6)	—	—	—
Neither agree nor disagree	143 (0)	60 (0)	—	—	—
Moderately agree	172 (–5.1)	119 (5.1)	—	—	—
Completely agree	29 (–5.8)	43 (5.8)	—	—	—
Self-reported training[b] (n=978)			**<.001**	**χ^2_4=94.6**	**.31**
Completely disagree	301 (7.4)	54 (–7.4)	—	—	—
Moderately disagree	188 (0.1)	78 (–0.1)	—	—	—

Characteristic	Previous experience of use[a]		Chi-square tests		
	No	Yes	*P* value	Chi-square	Cramér V Φc
Neither agree nor disagree	115 (–0.8)	54 (0.8)	—	—	—
Moderately agree	65 (–6.2)	71 (6.2)	—	—	—
Completely agree	20 (–5.2)	32 (5.2)	—	—	—
Recommendation[c] (n=1077)			**<.001**	**χ^2_4=37.3**	**.19**
No	649 (6.1)	222 (–6.1)	—	—	—
Yes	109 (–6.1)	97 (6.1)	—	—	—
Attitudes (n=972)			**<.001**	**χ^2_3=38.4**	**.20**
≤36.02	190 (2.9)	54 (–2.9)	—	—	—
36.03-46.20	185 (2.0)	60 (–2.0)	—	—	—
46.21-53.48	177 (1.2)	64 (–1.2)	—	—	—
≥53.49	133 (–6.1)	109 (6.1)	—	—	—

[a]Adjusted standardized residual frequencies appear in parentheses after observed group frequencies. Original wording: I am familiar with the concept of providing psychological support via the internet. Rated on a 5-point scale: 1=completely disagree to 5=completely agree.

[b]Original wording: I believe to have the necessary training to provide psychological support via the internet. Rated on a 5-point scale: 1=completely disagree to 5=completely agree.

[c]Original wording: Have you ever recommended the use of internet-based psychological support or other online resources in order to improve a client's health status?

The association of attitudes toward internet interventions with demographic and background factors, knowledge, training, recommendation, and future use was also assessed via chi-square analyses (see Table 5). These tests revealed a significant association between attitudes of respondents and self-reported knowledge, self-reported training, previous experience of use, recommendation, and future use of internet interventions. No significant associations were found between attitudes of respondents and demographic or background factors such as age, theoretical orientation, or professional experience.

Findings in these analyses primarily reflect the fact that psychologists without any knowledge, training, or previous experience using internet interventions are more likely to present more negative attitudes toward these interventions than expected. Conversely, psychologists reporting moderate to high knowledge, adequate training, and prior experience on the implementation of internet interventions were more prone to present favorable attitudes toward these interventions.

Additionally, participants having more positive attitudes toward internet interventions had a higher probability of recommending internet interventions and online resources to improve the health status of their clients and considering using such interventions in the future. Opposingly, participants presenting more negative attitudes toward internet interventions were less likely to recommend or contemplate using such interventions in the future. No demographic or background factors were significantly associated with attitudes toward internet interventions in this study.

Table 5. Factors associated with attitudes toward internet interventions.

Characteristic	Attitudes				Chi-square tests		
	≤36.02	36.03-46.20	46.21-53.48	≥53.49	*P* value	Chi-square	Cramér V
Self-reported knowledge[a] (n=972)					**<.001**	**χ^2_{12}=109.74**	**.19**
Completely disagree	63 (4.6)	45 (0.9)	32 (–1.5)	20 (–4.0)			
Moderately disagree	68 (0.8)	83 (3.3)	60 (–0.4)	40 (–3.8)			
Neither agree nor disagree	56 (1.0)	50 (–0.1)	56 (1.1)	39 (–2.0)			
Moderately agree	49 (–3.8)	57 (–2.6)	81 (1.5)	102 (4.9)			
Completely agree	8 (–2.8)	10 (–2.2)	12 (–1.6)	41 (6.6)			
Self-reported training[b] (n=972)					**<.001**	**χ^2_{12}=64.70**	**.15**
Completely disagree	109 (3.1)	103 (2.2)	86 (–0.2)	55 (–5.1)			
Moderately disagree	64 (–0.5)	64 (–0.5)	77 (1.8)	61 (–0.9)			
Neither agree nor disagree	41 (–0.1)	43 (0.2)	41 (0)	41 (–0.1)			
Moderately agree	22 (–2.6)	25 (–2.0)	27 (–1.4)	62 (6.0)			
Completely agree	8 (–1.6)	10 (–0.9)	10 (–0.9)	23 (3.4)			
Recommendation[c] (n=972)					**<.001**	**χ^2_3=42.12**	**.21**
No	223 (4.8)	204 (1.1)	194 (–0.2)	166 (–5.7)			
Yes	21 (–4.8)	41 (–1.1)	47 (0.2)	76 (5.7)			
Future use[d] (n=685)					**<.001**	**χ^2_3=123.14**	**.42**
No	132 (10.5)	65 (–0.1)	44 (–4.2)	20 (–6.1)			
Yes	58 (–10.5)	120 (1.0)	133 (4.2)	113 (6.1)			

[a]Adjusted standardized residual frequencies appear in parentheses after observed group frequencies. Original wording: I am familiar with the concept of providing psychological support via the internet. Rated on a 5-point scale: 1=completely disagree to 5=completely agree.

[b]Original wording: I believe to have the necessary training to provide psychological support via the internet. Rated on a 5-point scale: 1=completely disagree to 5=completely agree.

[c]Original wording: Have you ever recommended the use of internet-based psychological support or other online resources in order to improve a client's health status?

[d]Original wording: Do you expect to use the internet or the telephone to provide psychological support in the future?

Discussion

Principal Findings

The aim of this study was to explore Portuguese psychologist knowledge, training, use, and attitudes toward internet interventions, investigate perceived advantages and limitations of such interventions, identify potential drivers and barriers impacting implementation, and study potential factors associated to use and attitudes toward internet interventions.

Results showed that most psychologists were not familiar with internet interventions and had no prior experience using digital technology in the provision of psychological support. Only a minority reported having the necessary training to work in the field. Nevertheless, more than half of nonusers contemplated using it in the future, mainly as blended and pharmacotherapy complementary interventions rather than stand-alone internet interventions. From those who had prior experience implementing such programs, the majority were clinical and health psychologists who used telephone, email, and SMS services as a complement to face-to-face interventions with the purpose of increasing access to information and psychological care when treating mental health disorders such as anxiety or depression. Guided and unguided internet interventions were rarely used in this context. These results are in line with previous studies [15,34,46] that showed a higher acceptance of blended interventions when compared with stand-alone internet interventions but contrast with the reality of countries such as Australia [26], the United Kingdom, and Sweden [34], where the use of internet interventions is widely disseminated. As conceptualized by Topooco et al [34], Portugal, in this domain, may be included in the learners category, since the current experience and practice of e-mental health in the country is very limited.

Although accessibility, convenience, and cost effectiveness are considered important advantages of internet interventions by Portuguese psychologists, their attitudes toward such interventions tend to range from slightly negative to neutral, and a guarded stance is adopted when analyzing the topic.

Similar findings were reported by Neves et al [24] in a study assessing the impact of evidence-based practice on the attitudes of Portuguese psychologists toward internet interventions. Perceived barriers and limitations associated to internet interventions implementation may contribute to this predisposition and partly explain the low uptake of these interventions in Portugal.

According to participants in this study, the main barriers to overcome in the implementation of internet interventions were related to limitations on the conceptual comprehension and implementation of self-help techniques by clients, insufficient scientific evidence on the efficacy and cost effectiveness of internet interventions, and difficulties in the adaptation of treatment protocols to the digital format. Although these may be in fact challenges to overcome in some domains, the high number of publications attesting to the efficacy and cost effectiveness of internet interventions based on established treatment protocols and promoting the use of self-help techniques by clients [47-49] refutes these misconceptions. Other important obstacles identified by psychologists participating in this study pertained to patient ICT illiteracy and low adherence. Considering that in 2018 [50], 79% of Portuguese households had access to the internet, 75% of residents in the country aged 16 to 74 years reported using the internet in the previous year, mainly via smartphones, and 67% and 80% used apps and authentication procedures, respectively, ICT illiteracy may still be a barrier in some users over age 55 years and in extremely remote regions, but the necessary conditions to implement internet interventions successfully in Portugal are already in place.

As in previous studies [15,21,27-29], security, confidentiality, and ethical concerns were other important obstacles identified by participants in this research. On one hand, the fact that in 2018 alone several social media companies such as Facebook and Google reported data breaches [51] compromising the personal information of millions of users around the world may contribute to an atmosphere of insecurity and suspicion toward information systems security. On the other hand, the fact that until May 2019, no guidelines for the practice of e-mental health had been published by OPP [52] providing practical and deontological orientation in this domain may have contributed to psychologist reluctance in using information and communication systems in their practice. As we have stated, the provision of psychological services via digital technology has idiosyncrasies and must comply with ethical and security requirements similar to those used in online banking, which may at the same time promote confidence and inhibit the use of internet interventions, depending on psychologist ICT literacy.

Another important aspect relating to the ethics and process of delivering psychological support via the internet that occasionally emerged in this research as potentially affecting implementation pertains to the deleterious effect internet interventions may have on psychological assessment, therapeutic alliance, and crisis management. Like in previous publications [19,28,53], we found that a significant proportion of Portuguese psychologists perceive remote psychological assessment processes as inaccurate, increasing the possibility of misdiagnosis, and unsuitable for crisis management. As reported by Vigerland et al [22] and Perle et al [21], internet intervention adoption by psychologists may depend on disorder type and tend to be rejected for more serious conditions, despite the growing evidence attesting its efficacy in severe disorders [54-56]. Moreover, a disbelief in the possibility of establishing an adequate therapeutic alliance via the internet was reported by approximately one-third of respondents in this study, corroborating the findings of Sucala et al [53]. In that study, clinicians reported less confidence in their skills to develop alliance in e-therapy than in face-to-face therapy, mainly due to anticipated difficulties in reading patient emotions and conveying warmth and empathy in this environment. Although research on this topic is scarce, recent studies suggest therapeutic alliance in internet-based cognitive behavioral therapy is high [57,58] and has the potential of enhancing engagement and rapport in face-to-face psychotherapy when combined [59].

Last, negative attitudes presented both by patients and psychologists toward internet interventions were other important obstacles identified by participants in this research. The assessment of their attitudes exhibited in the context of this study confirmed this assertion. However, although only a few studies focused on the acceptability of internet interventions, the existing literature seems to point in the opposite direction [39,60] and suggests individuals with depression hold more positive attitudes toward such interventions than do psychotherapists [23]. In Portugal, a study focusing on Portuguese women's acceptance of e-mental health tools during the perinatal period reported good acceptance of internet interventions by this group. Nevertheless, more research focusing on different patient groups is necessary to adequately characterize the attitudes of patients with mental health disorders.

Regarding potential factors associated with Portuguese psychologist use of internet interventions, a significant association was found between previous experience of use and age, years of professional experience, work context, theoretical orientation, attitudes, knowledge, training, and recommendation of internet interventions. Unexpectedly, digital native psychologists (aged 30 years and younger) and psychologists with less than 5 years of professional experience were less likely to have used internet interventions in the past when compared with their middle-aged (aged 41 to 60 years) and more experienced colleagues (16 or more years), a finding that is not justified by a delay entering the labor market, since most of our sample was active and no significant differences were found regarding work status between the different age groups. Furthermore, considering that no significant associations were found between age and attitudes toward internet interventions, this finding might be justified by seasoned psychologists feeling more in control of the therapeutic process and therefore more lenient toward setting rules and more willing to use innovative tools in their practice. Work context also seems to impact internet intervention adoption. Psychologists working at the NHS and in private practices were more likely to include digital technology in the therapeutic process than psychologists working at public services, education/research institutions, and charities. The shortage of mental health professionals [61,62] working at

a universal, general, and tendentiously free NHS as is the case in Portugal may possibly burden the class and incentivize the use of creative solutions for patients' problems. As identified in the research by Venkatesh et al [33], performance expectancy (usefulness, effectiveness, enhancement of quality, diversity of care, and increase in productivity) and facilitating conditions seem to emerge as possible predictors of use. In private practice, the nature of the psychotherapeutic relationship and possible requests from long-term patients living in a globalized digital world probably impose such solutions. This aspect may also justify another surprising finding in this study. Contrasting with previous research [22-24], theoretical orientation was not significantly associated with attitudes toward internet interventions, but dynamically oriented therapists were more likely to have used internet interventions in the past than their colleagues with other theoretical stances. Besides the fact that psychodynamic interventions are typically longer and therefore possibly more challenged by the necessity of including alternative ways of communications in the process, the nonprescriptive nature of psychodynamic psychotherapy may turn dynamically oriented therapists less susceptible to fear of replacement and consequently more open to include digital communication tools in their practice. Although this study did not pursue this line of inquiry, an analysis of participant attitudes toward internet interventions revealed a significant association between previous use and psychologist predisposition toward such interventions, supporting the thesis that attitudes impact adoption. In this study, psychologists presenting more negative attitudes toward internet interventions were less likely to have prior experience using internet interventions when compared with psychologists holding more positive attitudes. Significant associations were also identified between previous use and attitudes toward internet interventions and self-reported knowledge, self-reported training, recommendation, and future use. Psychologists reporting ignorance on the subject and having no training in internet interventions were more likely to present more negative attitudes toward these interventions and have less experience in their implementation. Conversely, psychologists reporting moderate to high knowledge, adequate training, and prior experience in the implementation of these interventions were more prone to present favorable attitudes toward internet interventions, which confirms the findings of Whitfield and Williams [46] and Glueckauf et al [27] and identifies lack of knowledge and training as major barriers to overcome in this context. Naturally, psychologists with previous implementation experience and more positive attitudes toward internet interventions had a higher probability of recommending internet interventions to their clients and contemplating their use in the future.

Considering the prevalence of lifetime mental health disorders in Portugal is above 30% [17], the Portuguese mental health system is failing to comply with World Health Organization recommendations of providing better access and more integrated mental health care to the Portuguese population [63], and Portuguese psychologist attitudes toward evidence-based internet interventions are hindering implementation, mainly due to lack of knowledge and training, immediate corrective actions must be taken. Awareness and training initiatives should be promoted by psychologist associations and universities as an effective means of educating the class, changing professional and student perceptions [25,64,65], and increasing the implementation of ubiquitous strategies such as internet interventions to equitably respond to the mental health care system challenges and limitations. Additional research focusing on the Portuguese e-mental health ecosystem and addressing most prevalent mental health disorders in the country should inform and result from this process.

Limitations and Future Work

Several limitations must be considered when interpreting our findings. Despite ATIIS good psychometric properties, the fact that the two selected factors—positive and negative attitudes—only account for 44% of the variance explained suggests further research is necessary to understand what other factors might be attributable to psychologist attitudes toward internet interventions. Second, ATIIS online dissemination and the study sample self-selection might have introduced selection bias, limiting the generalizability of the obtained results. ICT illiterate psychologists as well as those presenting more negative attitudes toward internet interventions might not have participated in this study, lowering the response rate and biasing its results. Nevertheless, the study sample may be considered very large, and its demographic and background characteristics are similar to those published on the last census performed by OPP [44], indicating that participants in this study are probably representative of the class. Moreover, findings in our research are concordant with those reported by Neves et al [24], supporting external validity. Third, the exploratory cross-sectional design adopted in this investigation and the fact that no theoretical framework was used to present a structured representation of factors influencing adoption may also be considered a limitation. Additional research adopting structured frameworks and resourcing to mixed-methods research should be performed in order to deeply explore adoption and attitude predictors. To this end, a complementary study, adopting a qualitative descriptive approach consisting of in-depth semistructured interviews with Portuguese psychologists, is being conducted by the research team. Finally, the fact that this study targeted only psychologists and not other stakeholders in the e-mental health ecosystem such as patients/service-users, other health care providers, government bodies, funding/insurance bodies, technical developers, and researchers fails to present a comprehensive picture of the current status and acceptability of e-mental health in Portugal. To address this limitation, the research team is conducting a mixed-methods research study to explore the attitudes of Portuguese breast cancer patients toward internet interventions. In that study, ATIIS was adapted to the target population, and its factor structure will be examined. Future research should characterize the attitudes of other e-mental health ecosystem stakeholders.

Conclusions

This study investigated the use and attitudes of Portuguese psychologists toward internet interventions and provided insight on the principal barriers hindering implementation in the country. Most Portuguese psychologists were not familiar with and had no training or prior experience using internet interventions. A slightly negative/neutral attitude toward internet

interventions was captured, indicating that Portuguese psychologists are cautious toward these interventions and show greater acceptability toward blended treatment interventions compared with stand-alone internet interventions. Lack of knowledge and training are likely the main barriers to overcome for successful implementation and underline the need for awareness and training initiatives focusing not only on internet intervention efficacy and cost effectiveness but also on the practical, relational, technological, ethical, and regulatory requirements this treatment modality entails.

Acknowledgments

The authors wish to thank OPP for advertising the study to its members. The authors acknowledge the assistance of George Vlaescu in the acquisition of the data. CMS was funded by the Erasmus+ Program of the European Union.

Authors' Contributions

This study was conceptualized and designed by CMS, EW, RS, and GA. CMS acquired the data, analyzed and interpreted the data, and wrote the manuscript. EW, RS, and GA revised the article for important intellectual content. All authors equally contributed to this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Attitudes Toward Internet Interventions Survey factor analysis (rotated component matrix).
[DOCX File , 15 KB - mental_v7i4e16817_app1.docx]

References

1. World Health Organization. WHO guideline: recommendations on digital interventions for health system strengthening: research considerations URL: http://www.who.int/iris/handle/10665/311978 [accessed 2020-02-11]
2. 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]
3. 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]
4. Olthuis JV, Watt MC, Bailey K, Hayden JA, Stewart SH. Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. Cochrane Database Syst Rev 2015;3:CD011565. [doi: 10.1002/14651858.CD011565] [Medline: 25742186]
5. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. J Technol Hum Serv 2008 Jul 03;26(2-4):109-160. [doi: 10.1080/15228830802094429]
6. Hadjiconstantinou M, Byrne J, Bodicoat DH, Robertson N, Eborall H, Khunti K, et al. Do web-based interventions improve well-being in type 2 diabetes? A systematic review and meta-analysis. J Med Internet Res 2016 Oct 21;18(10):e270 [FREE Full text] [doi: 10.2196/jmir.5991] [Medline: 27769955]
7. Agboola SO, Ju W, Elfiky A, Kvedar JC, Jethwani K. The effect of technology-based interventions on pain, depression, and quality of life in patients with cancer: a systematic review of randomized controlled trials. J Med Internet Res 2015;17(3):e65 [FREE Full text] [doi: 10.2196/jmir.4009] [Medline: 25793945]
8. Kuijpers W, Groen WG, Aaronson NK. A systematic review of web-based interventions for patient empowerment and physical activity in chronic diseases: relevance for cancer survivors. J Med Internet Res 2013;15(2):e37 [FREE Full text] [doi: 10.2196/jmir.2281] [Medline: 23425685]
9. Loucas CE, Fairburn CG, Whittington C, Pennant ME, Stockton S, Kendall T. E-therapy in the treatment and prevention of eating disorders: a systematic review and meta-analysis. Behav Res Ther 2014 Dec;63:122-131 [FREE Full text] [doi: 10.1016/j.brat.2014.09.011] [Medline: 25461787]
10. Wolvers MD, Bruggeman-Everts FZ, Van der Lee ML, Van de Schoot R, Vollenbroek-Hutten MM. Effectiveness, mediators, and effect predictors of internet interventions for chronic cancer-related fatigue: the design and an analysis plan of a 3-armed randomized controlled trial. JMIR Res Protoc 2015 Jun 23;4(2):e77 [FREE Full text] [doi: 10.2196/resprot.4363] [Medline: 26104114]
11. Schröder J, Berger T, Westermann S, Klein JP, Moritz S. Internet interventions for depression: new developments. Dialogues Clin Neurosci 2016 Dec;18(2):203-212 [FREE Full text] [Medline: 27489460]
12. Carlbring P, Andersson G. Internet and psychological treatment. How well can they be combined? Comput Hum Behav 2006 May;22(3):545-553. [doi: 10.1016/j.chb.2004.10.009]

13. 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]
14. Andersson G. Internet-delivered psychological treatments. Annu Rev Clin Psychol 2016;12:157-179. [doi: 10.1146/annurev-clinpsy-021815-093006] [Medline: 26652054]
15. Feijt MA, de Kort YA, Bongers IM, IJsselsteijn WA. Perceived drivers and barriers to the adoption of emental health by psychologists: the construction of the levels of adoption of emental health model. J Med Internet Res 2018 Dec 24;20(4):e153 [FREE Full text] [doi: 10.2196/jmir.9485] [Medline: 29691215]
16. Tonn P, Reuter SC, Kuchler I, Reinke B, Hinkelmann L, Stöckigt S, et al. Development of a questionnaire to measure the attitudes of laypeople, physicians, and psychotherapists toward telemedicine in mental health. JMIR Ment Health 2017 Oct 03;4(4):e39 [FREE Full text] [doi: 10.2196/mental.6802] [Medline: 28974485]
17. Caldas-de-Almeida J, Xavier M. [Epidemiological study of mental health]. Lisbon: Universidade Nova de Lisboa; 2013.
18. Schuster R, Pokorny R, Berger T, Topooco N, Laireiter A. The advantages and disadvantages of online and blended therapy: survey study amongst licensed psychotherapists in Austria. J Med Internet Res 2018 Dec 18;20(12):e11007 [FREE Full text] [doi: 10.2196/11007] [Medline: 30563817]
19. Mora L, Nevid J, Chaplin W. Psychologist treatment recommendations for Internet-based therapeutic interventions. Comput Hum Behav 2008 Sep;24(6):3052-3062. [doi: 10.1016/j.chb.2008.05.011]
20. Wangberg SC, Gammon D, Spitznogle K. In the eyes of the beholder: exploring psychologists' attitudes towards and use of e-therapy in Norway. Cyberpsychol Behav 2007 Jun;10(3):418-423. [doi: 10.1089/cpb.2006.9937] [Medline: 17594266]
21. Perle JG, Langsam LC, Randel A, Lutchman S, Levine AB, Odland AP, et al. Attitudes toward psychological telehealth: current and future clinical psychologists' opinions of internet-based interventions. J Clin Psychol 2013 Jan;69(1):100-113. [doi: 10.1002/jclp.21912] [Medline: 22975897]
22. Vigerland S, Ljótsson B, Bergdahl Gustafsson F, Hagert S, Thulin U, Andersson G, et al. Attitudes towards the use of computerized cognitive behavior therapy (cCBT) with children and adolescents: a survey among Swedish mental health professionals. Internet Interventions 2014 Jul;1(3):111-117 [FREE Full text] [doi: 10.1016/j.invent.2014.06.002]
23. Schröder J, Berger T, Meyer B, Lutz W, Hautzinger M, Späth C, et al. Attitudes towards internet interventions among psychotherapists and individuals with mild to moderate depression symptoms. Cogn Ther Res 2017 Apr 22;41(5):745-756. [doi: 10.1007/s10608-017-9850-0]
24. Neves FRP. O efeito da prática profissional baseada na evidência nas atitudes dos psicólogos face à intervenção psicológica eletrónica [Dissertation]. 2017. URL: http://recil.grupolusofona.pt/handle/10437/8757 [accessed 2019-10-28]
25. Simms DC, Gibson K, O'Donnell S. To use or not to use: clinicians' perceptions of telemental health. Canadian Psychology/Psychologie canadienne 2011 Feb;52(1):41-51. [doi: 10.1037/a0022275]
26. Bruno R, Abbott JM. Australian health professionals' attitudes toward and frequency of use of internet supported psychological interventions. Int J Mental Health 2015 Jun 15;44(1-2):107-123. [doi: 10.1080/00207411.2015.1009784]
27. Glueckauf RL, Maheu MM, Drude KP, Wells BA, Wang Y, Gustafson DJ, et al. Survey of psychologists' telebehavioral health practices: technology use, ethical issues, and training needs. Prof Psychol Res Pract 2018 Jun;49(3):205-219. [doi: 10.1037/pro0000188]
28. Evans D. South African psychologists' use of the Internet in their practices. South African J Psychol 2014;44(2):162-169.
29. Cipolletta S, Mocellin D. Online counseling: an exploratory survey of Italian psychologists' attitudes towards new ways of interaction. Psychother Res 2018 Nov;28(6):909-924. [doi: 10.1080/10503307.2016.1259533] [Medline: 28068875]
30. Lazuras L, Dokou A. Mental health professionals' acceptance of online counseling. Technol Soc 2016 Feb;44:10-14. [doi: 10.1016/j.techsoc.2015.11.002]
31. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
32. Schulze N, Reuter SC, Kuchler I, Reinke B, Hinkelmann L, Stöckigt S, et al. Differences in attitudes toward online interventions in psychiatry and psychotherapy between health care professionals and nonprofessionals: a survey. Telemed J E Health 2018 Nov 09. [doi: 10.1089/tmj.2018.0225] [Medline: 30412450]
33. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
34. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, et al. Attitudes towards digital treatment for depression: a European stakeholder survey. Internet Interv 2017 Jun;8:1-9. [doi: 10.1016/j.invent.2017.01.001]
35. Mendes-Santos C, Weiderpass E, Santana R, Andersson G. A guided internet-delivered individually-tailored ACT-influenced cognitive behavioural intervention to improve psychosocial outcomes in breast cancer survivors (iNNOVBC): study protocol. Internet Interv 2019 Sep;17:100236 [FREE Full text] [doi: 10.1016/j.invent.2019.01.004] [Medline: 30949435]
36. LimeSurvey: An Open Source Survey Tool. Hamburg: Limesurvey GmbH; 2019.
37. Schröder J, Sautier L, Kriston L, Berger T, Meyer B, Späth C, et al. Development of a questionnaire measuring Attitudes towards Psychological Online Interventions-the APOI. J Affect Disord 2015 Nov 15;187:136-141. [doi: 10.1016/j.jad.2015.08.044] [Medline: 26331687]

38. Jansen F, van Uden-Kraan UCF, van Zwieten ZV, Witte BI, Verdonck-de Leeuw IM. Cancer survivors' perceived need for supportive care and their attitude towards self-management and eHealth. Support Care Cancer 2015 Jun;23(6):1679-1688. [doi: 10.1007/s00520-014-2514-7] [Medline: 25424520]
39. Yao X, Li Z, Arthur D, Hu L, Cheng G. The feasibility of an internet-based intervention for Chinese people with mental illness: a survey of willingness and attitude. Int J Nurs Sci 2014 Mar;1(1):28-33. [doi: 10.1016/j.ijnss.2014.02.008]
40. Kivi M, Eriksson MC, Hange D, Petersson E, Björkelund C, Johansson B. Experiences and attitudes of primary care therapists in the implementation and use of internet-based treatment in Swedish primary care settings. Internet Interv 2015 Sep;2(3):248-256. [doi: 10.1016/j.invent.2015.06.001]
41. Ajzen I. Nature and operation of attitudes. Annu Rev Psychol 2001;52:27-58. [doi: 10.1146/annurev.psych.52.1.27] [Medline: 11148298]
42. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
43. Carlson JE, Stevens J. Applied multivariate statistics for the social sciences. J Educ Stat 1988;13(4):368. [doi: 10.2307/1164712]
44. Os números da Ordem: Censo dos Membros Efectivos da Ordem dos Psicólogos Portugueses. 2014 Oct 07. URL: https://www.ordemdospsicologos.pt/pt/noticia/1246 [accessed 2019-10-28]
45. George D, Mallery P. SPSS for Windows Step by Step: A Simple Guide and Reference, 11.0 Update. Boston: Allyn and Bacon; 2003.
46. Whitfield G, Williams C. If the evidence is so good, why doesn't anyone use them? A national survey of the use of computerized cognitive behaviour therapy. Behav Cognit Psychother 1999;32(1):57-65. [doi: 10.1017/S1352465804001031]
47. Andersson G. Internet interventions: past, present and future. Internet Interv 2018 Jun;12:181-188. [doi: 10.1016/j.invent.2018.03.008]
48. 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]
49. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
50. [Information and knowledge society: Survey on the use of information and communication technologies by families]. Lisbon: Instituto Nacional de Estatística; 2018.
51. The 21 scariest data breaches of 2018. 2018 Dec 30. URL: https://www.businessinsider.com/data-hacks-breaches-biggest-of-2018-2018-12 [accessed 2020-02-13]
52. Carvalho R, Fonseca A, Dores A, Mendes-Santos C, Batista J, Salgado B, et al. Ordem dos Psicólogos Portugueses. 2015 May 31. Linhas de Orientação para a Prestação de Serviços de Psicologia Mediados por Tecnologias da Informação e da Comunicação URL: https://www.ordemdospsicologos.pt/ficheiros/documentos/guidelines_opp_psicologia_ehealth.pdf [accessed 2019-10-28]
53. Sucala M, Schnur JB, Brackman EH, Constantino MJ, Montgomery GH. Clinicians' attitudes toward therapeutic alliance in E-therapy. J Gen Psychol 2013;140(4):282-293 [FREE Full text] [doi: 10.1080/00221309.2013.830590] [Medline: 24837821]
54. Löbner M, Pabst A, Stein J, Dorow M, Matschinger H, Luppa M, et al. Computerized cognitive behavior therapy for patients with mild to moderately severe depression in primary care: a pragmatic cluster randomized controlled trial (@ktiv). J Affect Disord 2018 Oct 01;238:317-326. [doi: 10.1016/j.jad.2018.06.008] [Medline: 29902736]
55. Meyer B, Bierbrodt J, Schröder J, Berger T, Beevers CG, Weiss M, et al. Effects of an Internet intervention (Deprexis) on severe depression symptoms: randomized controlled trial. Internet Interv 2015 Mar;2(1):48-59. [doi: 10.1016/j.invent.2014.12.003]
56. Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. BMJ 2013;346:f540 [FREE Full text] [Medline: 23444423]
57. Pihlaja S, Stenberg J, Joutsenniemi K, Mehik H, Ritola V, Joffe G. Therapeutic alliance in guided internet therapy programs for depression and anxiety disorders: a systematic review. Internet Interv 2018 Mar;11:1-10 [FREE Full text] [doi: 10.1016/j.invent.2017.11.005] [Medline: 30135754]
58. Berger T. The therapeutic alliance in internet interventions: a narrative review and suggestions for future research. Psychother Res 2017 Sep;27(5):511-524. [doi: 10.1080/10503307.2015.1119908] [Medline: 26732852]
59. Richards P, Simpson S, Bastiampillai T, Pietrabissa G, Castelnuovo G. The impact of technology on therapeutic alliance and engagement in psychotherapy: the therapist's perspective. Clin Psychol 2016 Aug 05;22(2):171-181. [doi: 10.1111/cp.12102]
60. Fonseca A, Gorayeb R, Canavarro MC. Women's use of online resources and acceptance of e-mental health tools during the perinatal period. Int J Med Inform 2016 Oct;94:228-236. [doi: 10.1016/j.ijmedinf.2016.07.016] [Medline: 27573331]

61. World Health Organization. 2011 Feb 16. WHO Mission to assess the progress of the mental health reforms in Portugal URL: https://tinyurl.com/tqcb6uc [accessed 2020-02-11]
62. Entidade Reguladora da Saúde. 2015 Sep. Acesso e qualidade nos cuidados de saúde mental URL: https://www.ers.pt/uploads/writer_file/document/1500/Estudo_Saude_Mental__versao_publicar__v.2.pdf [accessed 2019-10-28]
63. Perelman J, Chaves P, de Almeida JMC, Matias MA. Reforming the Portuguese mental health system: an incentive-based approach. Int J Ment Health Syst 2018;12:25 [FREE Full text] [doi: 10.1186/s13033-018-0204-4] [Medline: 29853991]
64. Donovan CL, Poole C, Boyes N, Redgate J, March S. Australian mental health worker attitudes towards cCBT: What is the role of knowledge? Are there differences? Can we change them? Internet Interv 2015 Nov;2(4):372-381. [doi: 10.1016/j.invent.2015.09.001]
65. van der Vaart R, Atema V, Evers AWM. Guided online self-management interventions in primary care: a survey on use, facilitators, and barriers. BMC Fam Pract 2016 Mar 09;17:27 [FREE Full text] [doi: 10.1186/s12875-016-0424-0] [Medline: 26961547]

Abbreviations

ATIIS: Attitudes Toward Internet Interventions Survey
EFA: exploratory factor analysis
EPI: electronic psychological interventions
ICT: information and communication technologies
iNNOVBC: A Guided Internet-Delivered Individually Tailored ACT-Influenced Cognitive Behavioral Intervention to Improve Psychosocial Outcomes in Breast Cancer Survivors
KMO: Kaiser-Meyer-Olkin test
NHS: National Health Service
OPP: Portuguese Psychologists Association (Ordem dos Psicólogos Portugueses)
SMS: short message service

Edited by G Eysenbach; submitted 28.10.19; peer-reviewed by M Herdeiro, P Navarro; comments to author 24.11.19; revised version received 10.12.19; accepted 26.01.20; published 06.04.20.

Please cite as:
Mendes-Santos C, Weiderpass E, Santana R, Andersson G
Portuguese Psychologists' Attitudes Toward Internet Interventions: Exploratory Cross-Sectional Study
JMIR Ment Health 2020;7(4):e16817
URL: https://mental.jmir.org/2020/4/e16817
doi:10.2196/16817
PMID:32250273

Letter to the Editor

Comment on "Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow": A UK Perspective

Pauline Whelan[1,2], BA, MSc, PhD; Charlotte Stockton-Powdrell[1], MA; Jenni Jardine[2], BSc, ClinPsy; John Sainsbury[3], MSc

[1]Centre for Health Informatics, University of Manchester, Manchester, United Kingdom

[2]CAMHS.Digital Research Unit, Greater Manchester Mental Health NHS Foundation Trust, Manchester, United Kingdom

[3]Greater Manchester Mental Health NHS Foundation Trust, Manchester, United Kingdom

Corresponding Author:
Pauline Whelan, BA, MSc, PhD
Centre for Health Informatics
University of Manchester
Vaughan House
Manchester, M13 9PL
United Kingdom
Phone: 44 7554227474
Email: pauline.whelan@manchester.ac.uk

Related Article:

Comment on: https://mental.jmir.org/2020/3/e18848/

(JMIR Ment Health 2020;7(4):e19547) doi:10.2196/19547

KEYWORDS

digital mental health; digital psychiatry; COVID-19; mhealth; mobile apps; learning health system

This letter is in response to the article published by Torous et al [1], which highlighted the potential of digital mental health in improving the accessibility and quality of mental health service provision during and beyond the coronavirus disease (COVID-19) pandemic. We write from a UK perspective to add to the North American lens of Torous et al. We contribute our personal and multidisciplinary insights as members of a digital health software team at a UK research University (PW and CS-P), an innovation manager of a mental healthcare provider in England (JS), a digital mental health research unit for young people (PW and JJ), and directors of a digital mental health Community Interest Company (PW and CS-P).

We agree that robust evaluation of mental health apps is important, particularly when selecting an appropriate app during emergency conditions. In the United Kingdom, a national Apps Library provides a central point for accessing trusted apps. In response to the COVID-19 pandemic, a dedicated list of useful apps has been established by ORCHA (Organisation for the Review of Care and Health Applications), an organization that advises the National Health Service on the safety and efficacy of health apps. Apps recommended on the national Apps Library and by ORCHA have been through an independent review and evaluation process and can guide health professionals and patients to trusted apps to support mental health problems including anxiety and depression. The evaluation process covers data security, clinical evidence, and user experience. More broadly, the NICE (National Institute for Health and Care Excellence) recommends that digital health technologies define a set of national UK evidence standards to guide development and evaluation of digital mental health systems [2].

Informal feedback from clinicians implementing rapidly introduced digital innovations within our mental health services has emphasized the need for ongoing evaluation. We see value in creating digital learning health systems [3] to support iterative quality improvement and have aimed to do this across digital innovation projects. Situating planned changes within the NASSS (Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability) framework [4] has helped us assess how digital health innovations can be safely and sustainably embedded in care pathways. We value digital training for mental health professionals, a view reflected in the recent UK Topol review consultation [5].

Like Torous et al, we are concerned about how digital technologies may exacerbate health inequalities. However, our experience during COVID-19 of moving a face-to-face young people's digital mental health research group to an online

videoconference has highlighted how digital solutions can overcome pre-existing (or previously invisible) barriers to participation. Our online group meetings make the group more accessible for certain members due to reduced travel, time required, personal preferences, or specific mental health conditions, which had made face-to-face group time difficult. Similarly, when face-to-face contact is impossible for physical distancing reasons during COVID-19, digital solutions that can provide remote support fill a critical gap. We believe that a more equitable distribution of digital resources and adequate digital literacy provision will promote a healthier digital experience for all.

Conflicts of Interest

PW and CS-P are directors of Affigo CIC, a digital mental health community interest company.

References

1. 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]
2. NICE: National Institute for HealthCare Excellence. 2019. Evidence standards framework for digital health technologies URL: https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies [accessed 2019-01-24] [WebCite Cache ID 75fELArTW]
3. Friedman CP, Allee NJ, Delaney BC, Flynn AJ, Silverstein JC, Sullivan K, et al. The science of Learning Health Systems: Foundations for a new journal. Learn Health Syst 2017 Jan;1(1):e10020 [FREE Full text] [doi: 10.1002/lrh2.10020] [Medline: 31245555]
4. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
5. Foley T, Woollard J. NHS. 2019. The digital future of mental healthcare and its workforce: a report on a mental health stakeholder engagement to inform the Topol Review URL: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-Mental-health-paper.pdf [accessed 2020-04-24]

Abbreviations

COVID-19: coronavirus disease
NASSS: Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability
NICE: National Institute for Health and Care Excellence
ORCHA: Organisation for the Review of Care and Health Applications

Edited by J Torous; submitted 22.04.20; this is a non–peer-reviewed article;accepted 22.04.20; published 27.04.20.

Please cite as:
Whelan P, Stockton-Powdrell C, Jardine J, Sainsbury J
Comment on "Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow": A UK Perspective
JMIR Ment Health 2020;7(4):e19547
URL: http://mental.jmir.org/2020/4/e19547/
doi:10.2196/19547
PMID:32330113

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org