Feasibility, Acceptability, and Preliminary Efficacy of a Smartphone App–Led Cognitive Behavioral Therapy for Depression Under Therapist Supervision: Open Trial

Sabine Wilhelm¹, PhD; Emily E Bernstein¹, PhD; Kate H Bentley¹, PhD; Ivar Snorrason¹, PhD; Susanne S Hoeppner¹, PhD; Dalton Klare¹, MA; Jennifer L Greenberg¹, PsyD; Hilary Weingarden¹, PhD; Thomas H McCoy¹, MD; Oliver Harrison², MD

¹Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States
²Koa Health, London, United Kingdom

Corresponding Author:

Sabine Wilhelm, PhD Massachusetts General Hospital, Harvard Medical School 185 Cambridge Street Suite 2000 Boston, MA, 02114 United States Phone: 1 617 724 6146 Email: swilhelm@mgh.harvard.edu

Abstract

Background: Major depressive disorder affects approximately 1 in 5 adults during their lifetime and is the leading cause of disability worldwide. Yet, a minority receive adequate treatment due to person-level (eg, geographical distance to providers) and systems-level (eg, shortage of trained providers) barriers. Digital tools could improve this treatment gap by reducing the time and frequency of therapy sessions needed for effective treatment through the provision of flexible, automated support.

Objective: This study aimed to examine the feasibility, acceptability, and preliminary clinical effect of Mindset for Depression, a deployment-ready 8-week smartphone-based cognitive behavioral therapy (CBT) supported by brief teletherapy appointments with a therapist.

Methods: This 8-week, single-arm open trial tested the Mindset for Depression app when combined with 8 brief (16-25 minutes) video conferencing visits with a licensed doctoral-level CBT therapist (n=28 participants). The app offers flexible, accessible psychoeducation, CBT skills practice, and support to patients as well as clinician guidance to promote sustained engagement, monitor safety, and tailor treatment to individual patient needs. To increase accessibility and thus generalizability, all study procedures were conducted remotely. Feasibility and acceptability were assessed via attrition, patient expectations and feedback, and treatment utilization. The primary clinical outcome measure was the clinician-rated Hamilton Depression Rating Scale, administered at pretreatment, midpoint, and posttreatment. Secondary measures of functional impairment and quality of life as well as maintenance of gains (3-month follow-up) were also collected.

Results: Treatment credibility (week 4), expectancy (week 4), and satisfaction (week 8) were moderate to high, and attrition was low (n=2, 7%). Participants self-reported using the app or practicing (either on or off the app) the CBT skills taught in the app for a median of 50 (IQR 30-60; week 4) or 60 (IQR 30-90; week 8) minutes per week; participants accessed the app on an average 36.8 (SD 10.0) days and completed a median of 7 of 8 (IQR 6-8) steps by the week 8 assessment. The app was rated positively across domains of engagement, functionality, aesthetics, and information. Participants' depression severity scores decreased from an average Hamilton Depression Rating Scale score indicating moderate depression (mean 19.1, SD 5.0) at baseline to a week 8 mean score indicating mild depression (mean 10.8, SD 6.1; d=1.47; P<.001). Improvement was also observed for functional impairment and quality of life. Gains were maintained at 3-month follow-up.

Conclusions: The results show that Mindset for Depression is a feasible and acceptable treatment option for individuals with major depressive disorder. This smartphone-led treatment holds promise to be an efficacious, scalable, and cost-effective treatment option. The next steps include testing Mindset for Depression in a fully powered randomized controlled trial and real-world clinical settings.

Trial Registration: ClinicalTrials.gov NCT05386329; https://clinicaltrials.gov/study/NCT05386329?term=NCT05386329

(JMIR Ment Health 2024;11:e53998) doi: 10.2196/53998

KEYWORDS

depressive disorder; depressive; depression; open trial; open trials; single arm; smartphone; cognitive behavioral therapy; cBT; psychotherapy; psychoeducation; digital health; mobile applications; mHealth; mobile health; app; apps; application; applications; psychiatry; psychiatric; feasibility; acceptability; usability; satisfaction; user experience; mental

Introduction

Major depressive disorder (MDD), characterized by hallmark symptoms of persistent depressed mood and loss of interest in activities [1], is highly prevalent. In 2020, an estimated 21 million adults in the United States were impacted (8.4% population prevalence) [2]. The rates of elevated depressive symptoms have continued to rise since the COVID-19 pandemic, now affecting nearly 1 in 3 adults [3]. Depression is the leading cause of disability worldwide [4] and is associated with economic costs exceeding US \$326.2 billion in the United States alone [5]. Despite the substantial personal and societal impact of MDD, a minority of individuals meeting diagnostic criteria-let alone those at risk or with subthreshold symptoms-receive care; even fewer receive minimally adequate treatment (estimates range from 3% in low- to middle-income countries and 23% in high-income countries [6]), such as cognitive behavioral therapy (CBT), the most widely studied and recommended psychotherapy [7]. At a systems level, the limited availability of trained clinicians is a substantial contributor to low treatment utilization [8]. There simply are not and will not be enough clinicians to meet current demands for mental health care. Additionally, many people do not seek treatment for depression due to obstacles such as geographic distance from care providers, high costs, and stigma [8,9]. Given the prevalence of MDD and the substantial treatment gaps that exist, there is a clear need for low barrier, more widely accessible, effective treatments for MDD.

The technology could bridge such gaps. Smartphone apps or other digital tools could reduce the time and frequency of sessions by supplementing clinician effort with automated, validated support that can be used flexibly between sessions [10,11]. However, standalone apps are not sufficient or appealing for many patients [12,13]. The majority of apps, even those that are grounded in empirically supported treatments, have high dropout rates, which limits their effectiveness [14-16]. The absence of concurrent human support is often cited as the major reason for nonadherence or nonengagement [13,17-19]. Some engagement is likely a minimum requirement for an app-based therapy to be effective; guidance from a trained provider should further mitigate issues of comprehension, personalization, problem-solving, and interference from comorbid or life concerns [20]. Equivalent effects of face-to-face CBT and internet-delivered CBT for depression have been found for treatments that are therapist guided, meaning patients are in contact with a therapist throughout treatment (eg, weekly sessions, check-in phone calls, asynchronous, messaging, and weekly feedback emails) [21-23]. Moreover, many users simply want access to a therapist and are less willing to engage in

self-directed digital treatments [24,25]. Thus, a digital service that combines mobile-based CBT with brief remote individual sessions with a clinician (ie, teletherapy monitoring) has the potential to greatly enhance the scalability of high-quality app-based treatment, particularly for moderately and severely ill patients while reducing clinician burden and cost [26-28].

The purpose of this study was to conduct an open trial to test the feasibility, acceptability, and efficacy of the Mindset for Depression app (a novel, smartphone-based CBT program) with brief video-conferencing appointments with a therapist. We hypothesized that the treatment would be feasible and acceptable. We also hypothesized that treatment would yield statistically significant reductions in depression symptom severity (primary clinical outcome) as well as improvements in functioning and quality of life (secondary clinical outcomes) from baseline to posttreatment (week 8). The treatment was tested for patients with moderate to severe depression: those who would typically be referred for one-on-one outpatient therapy [29].

Methods

Study Design

This open trial tested the Mindset for Depression app when combined with brief (16-25 minutes) video-conferencing visits with a CBT therapist over 8 weeks. The primary outcomes were feasibility, acceptability, and preliminary efficacy, as measured by change in depression symptom severity. To increase accessibility and thus generalizability, all study procedures were conducted remotely.

Ethical Considerations

The study was approved by the institutional review board of Massachusetts General Hospital (2020P001958). All participants provided informed consent prior to the initiation of study procedures and were given the ability to opt out at any point. Data were deidentified to protect participants' privacy. Participants were compensated US \$25 at mid-treatment, end of treatment, and 3-month follow-up assessments.

Participants

Eligible participants, recruited between May 2022 and February 2023, were at least 18 years old, living in Massachusetts, presenting with a current primary *Diagnostic and Statistical Manual of Mental Disorders: 5th Edition (DSM-5)* diagnosis of MDD, and experiencing at least moderately severe symptoms (Patient Health Questionnaire-9 [PHQ-9] score \geq 10). Participants taking psychotropic medication were on a stable dose for at least 2 months prior to enrollment and were asked to remain on the same stable dose throughout the study period.

Exclusion criteria included 4 or more prior sessions of CBT for depression (assessed via self-report and interview with an independent evaluator), current severe substance use disorder, lifetime bipolar disorder or psychosis, acute and active suicidal ideation as indicated by clinical judgment, a score ≥ 2 on the past month suicidal ideation subscale of the Columbia-Suicide Severity Rating Scale [30], concurrent psychological treatment, and inability to engage with treatment (eg, did not own a supported smartphone).

Procedure

Treatment

The Mindset for Depression app provides key CBT-derived content for adults with MDD and was designed to be used in conjunction with a therapist over 8 weeks. The duration of CBT trials typically ranges from 6 to 20 sessions [31]. Mindset for Depression was built in collaboration between researchers at the Massachusetts General Hospital and Koa Health. The app-based format allows participants to review CBT content and accompanying skills practice exercises at their convenience and own pace and with support from their therapist.

The app and clinician dashboard were developed through collaborative, user-centered design, integrating perspectives from clinicians (MDs and psychologists with expertise in MDD and CBT), digital health researchers, patients with MDD and experience with CBT and other therapies, engineers, and designers. Through this approach, the product being tested was deployment ready (eg, built on a commercial platform, able to be quickly scaled and professionally maintained to minimize technical difficulties, and ensure compliance with up-to-date privacy and security standards) and therefore well positioned to succeed outside of research studies [32].

CBT Modules

The app delivers content in 8 steps, corresponding to the 8 weeks of treatment. A summary of these steps is visualized in Table 1. Participants were also allowed access to the app during the 3-month follow-up period. Core CBT skills included across treatment include psychoeducation, cognitive restructuring and core beliefs, behavioral activation, mindfulness, and relapse prevention [33,34]. Step 1 comprises psychoeducation about MDD and the CBT model and background and skills practice for identifying and restructuring "thinking traps," or maladaptive automatic thoughts [35,36]. Step 2 focuses on the short- and long-term impact of withdrawal and avoidance on mood and provides a structure for recording daily activities and monitoring associated moods. Step 3 introduces behavioral activation and scheduling and provides guidance for identifying valued or new activities and setting specific, measurable, achievable, relevant, and time-bound (SMART) goals [37,38]. Activity scheduling and monitoring (ie, recording completed activities and associated mood ratings), with an emphasis on personal values and meaning, continue for the remainder of treatment. Step 4 introduces mindfulness (present-focused and nonjudgmental awareness) and offers a guided mindful breathing audio exercise [39]. Steps 5 and 6 provide users with additional mindfulness approaches, including grounding and letting go of unhelpful thoughts. Step 7 builds on prior cognitive skills and delves into the definition of core beliefs, their relationship to automatic thoughts and feelings, and strategies to identify and challenge them (eg, downward arrow technique and building self-esteem). Step 8 concludes with relapse prevention by helping users consolidate treatment skills, anticipate future challenges, and plan for continued practice and flexible use of skills. Example screenshots from the smartphone app are included in Figure 1.

Figure 1. Screenshots from the Mindset for Depression smartphone app. CBT: cognitive behavioral therapy; SMART: specific, measurable, achievable, relevant, and time-bound.



Table 1.	Summary	of steps in	n the Mindset	for Depression	program.
----------	---------	-------------	---------------	----------------	----------

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Psychoeducation	✓		·		ł		·	
Cognitive restructuring 🗸								
Relationship between behavior and mood		1						
Behavioral activation			\checkmark	\checkmark	\checkmark	1	1	1
Mindfulness				1	1	1		
Modifying core beliefs and building self- esteem							✓	
Relapse prevention								1

Therapists

Each participant was matched with a licensed doctoral-level therapist. Therapists were trained in and actively practicing CBT for MDD and provided with study-specific training in using the Mindset app and therapist dashboard prior to beginning the trial. To ensure proficiency, therapists were required to complete the Massachusetts General Hospital Psychiatry Academy CBT training course and pass (>90% correct) both the corresponding CBT knowledge test and an MDD knowledge test. Weekly supervision from the principal investigator (expert in CBT) was also provided. To ensure ongoing high-quality treatment, including that implementation fidelity targets were met and non-CBT techniques were absent, sessions were audio recorded, and 40 of the 224 planned sessions (18.9% of the 212 sessions ultimately conducted) were randomly selected and rated for competency and treatment adherence by an independent rater. Adherence raters were experienced in CBT for MDD and further trained and supervised. Core elements of each treatment session (5-6 items) were rated for adherence on a 7-point scale (1=not at all to 7=completely adherent) and then a global rating of adherence was assigned. The full adherence scale is included in Multimedia Appendix 1. Competence was rated on 12 aspects (32 items) of CBT for MDD (eg, positive outlook, knowledge, clear communication, empathy, flexibility, and empowering the patient). Each item was scored on a 5-point scale (1=not at all to 5=completely competent) and then a global rating of competence was assigned. Overall, adherence and competence were high, with 100% of all rated sessions evaluated as "completely" adherent and 100% of all rated sessions evaluated as "mostly" or "completely" competent.

Therapists offered each patient 8 video-conferencing appointments (16-25 minutes; via HIPAA [Health Insurance Portability and Accountability Act]–compliant video conference) to be conducted weekly. This duration of appointment corresponds to a clinician billing code (CPT-90832), helping to ensure that the reimbursement of clinician time would not become a barrier to scale-up following the research. As needed, because of a therapist's or patient's schedule, up to 2 sessions were able to be scheduled per week. Throughout the treatment, participants were able to communicate with their therapists between sessions through asynchronous in-app secure messaging. Sessions were meant to support a patient's progress through the app-led treatment. In this way, the model mimicked the "flipped classroom," a new pedagogical approach shown to improve student learning [40]. In a flipped classroom, students watch or read lectures and complete initial practice problems asynchronously, reserving valuable classroom time for active problem-solving with an instructor. As such, sessions were intended to monitor risk as needed, help participants set goals, enhance motivation, clarify and practice the skills learned via the Mindset app to best meet the patient's needs, brainstorm ideas for homework, and problem-solve treatment barriers that arose. Therapists were instructed to work within a CBT framework and not to introduce other treatment modalities. Such fidelity was monitored in weekly supervision, via therapist self-checks included within session records ("Did you use any of the following non-CBT techniques? [check all that apply]") and via adherence ratings (ie, the degree to which forbidden content was introduced). The therapist dashboard was a separate web-based portal wherein therapists could receive and respond to messages and track participant progress in the app.

Assessments

Assessments were conducted by master's or doctoral-level independent evaluators who were not involved in treatment, were complemented by participant self-report, and occurred at baseline, mid-treatment (week 4), end of treatment (week 8), and follow-up (3 months posttreatment). Evaluators completed training on all clinician-administered measures and were required to maintain high reliability (>0.75 intraclass correlation coefficient), with a gold standard expert rater; 18.9% (20/106) of randomly selected assessments were rated to prevent rater drift. Evaluators were not privy to participants' progress in treatment (eg, app content reviewed and session notes). Adverse events, life events, and changes in medication or outside treatment were surveyed at each assessment or when a patient reported to study staff.

Measures Descriptions

Baseline Diagnostic Assessment

The Mini International Neuropsychiatric Interview was used to establish eligibility and characterize the sample. It is a reliable, validated semistructured diagnostic assessment of *DSM-5* psychiatric disorders [41].

Feasibility and Acceptability

Participants completed the self-reported measures as follows. The Credibility/Expectancy Questionnaire (CEQ) [42], completed at baseline and week 4, is a 6-item, self-reported Likert-type questionnaire that assesses patients' judgments about

XSL•FO RenderX

the credibility of the treatment rationale and treatment expectancy. Items on both subscales are summed together for total outcome scores that can range from 3 to 27, where higher scores mean higher treatment credibility and higher outcome expectancy. We assessed the internal consistency of scales with coefficient omega (McDonald ω), given the heterogeneity of variances across scale items; coefficient ω can be interpreted in the same way as Cronbach α . The internal consistency of the credibility items in this sample ranged from ω =0.69 at baseline to ω =0.77 at week 4; for the expectancy items, internal consistency ranged from ω =0.82 at baseline to ω =0.96 at week 4. The Mobile Application Rating Scale User Version (uMARS) [43], administered at week 8, collects evaluations of mobile health apps. The 26 items assess participants' evaluations of engagement, functionality, aesthetics, information quality, app subjectivity quality, and perceived impact. Items are rated on differently worded 5-point Likert scales ranging from 1 (inadequate) to 5 (excellent). An overall app rating score can be calculated as the mean score of the first 4 subscales (engagement, functionality, aesthetics, and information quality; a range of 1-5), where higher scores indicate higher overall perceived app quality. In this sample, the internal consistency of the 4 subscales used in the overall mean scores were ω =0.76 for engagement, ω =0.83 for functionality, ω =0.72 for aesthetics, and ω =0.62 for information quality, with an overall item of ω=0.83. The Client consistency Satisfaction Questionnaire (CSQ) [44], completed at weeks 4 and 8, is an 8-item questionnaire assessing satisfaction with clinical services received. Each item uses a 4-point Likert scale. Items are summed for a total score ranging from 8 to 32, with higher scores indicating greater satisfaction. The internal consistency of the CSQ was ω =0.93 at week 4 and ω =0.89 at week 8. Treatment use was assessed with a single question: "On average, how much time (in minutes) do you spend using the app or practicing skills from the app in total, per week?" Answers were collected as the number of minutes in integer format, where more time spent on and off the app was interpreted as greater treatment use. In addition, app use data were collected automatically based on the actions participants completed in the app. Due to technical issues, 2 participants' app use data were inadvertently not recorded. The internal consistency values for the CEQ credibility subscale at baseline and the uMARS information quality subscale fell below 0.7 and are a noted limitation.

Clinician-Administered Measures

The primary measure of MDD symptom severity was the (clinician-rated) Hamilton Depression Rating Scale (HAM-D) [45]. Considered a gold standard means of assessing symptom severity in patients who are depressed, it contains 21 items that are rated on a mixture of 3- and 5-point Likert scales. The first 17 items are summed for the total score, which can range from 0 to 52. Higher scores indicate greater depression severity. The internal consistency of the HAM-D in this sample ranged from ω =0.79 at baseline to ω =0.93 at the 3-month follow-up (question 17 was necessarily omitted from internal consistency calculations due to the absence of variability in responses; all participants received a score of 0 for this "insight" item ["Acknowledges being depressed and ill"] at all assessment

points with the exception of 1 participant at the 3-month follow-up). To evaluate treatment response and remission, we used criteria of HAM-D score reductions of \geq 50% for treatment response, HAM-D score reductions of \geq 25% but < 50% for partial response, and HAM-D scores \leq 7 to indicate remission [46-48]. An expert rater reviewed 18.9% (20 of the 106 assessments that were ultimately completed) of HAM-D assessments. Inter-rater reliability was excellent (HAM-D: intraclass correlation coefficient (1,1)=0.91).

Self-Reported Measures

Participants completed the following secondary measures of symptoms and functioning at each assessment: (1) The Work and Social Adjustment Scale (WSAS) [49] is a 5-item, self-reported measure of impairment in occupational, social, and family domains. Items are measured on 9-point Likert scales ranging from 0 (no impairment at all) to 8 (very severe impairment). The items are summed for a total score ranging from 0 to 40, where higher scores mean higher functional impairment. The internal consistency of the WSAS ranged from ω =0.84 at baseline to ω =0.93 at the 3-month follow-up. (2) The Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form (Q-LES-Q-SF) [50] is a 16-item self-reported measure of subjective quality of life. Each question is rated on a 5-point Likert scale ranging from 1 (very poor) to 5 (very good). Questions 1-14 are then summed to a total score, and the total score is reported as a percentage maximum possible, such that the final percent score range is 0% to 100%; higher scores correspond to greater ratings of quality of life. The internal consistency of the Q-LES-Q-SF ranged from ω =0.84 at baseline to ω =0.90 at the 3-month follow-up. (3) The PHQ-9 [51] is a self-reported measure of the past week's depression severity. It includes 9 Likert scale items mapping onto DSM-5 symptom criteria and ranging from 0 (not at all) to 3 (every day). The internal consistency of the PHQ-9 ranged from ω =0.70 at baseline to ω =0.89 at week 8.

Data Analysis

Power Analysis

With 28 participants enrolled, we had > 80% power to detect pre- to posttreatment effect sizes of $d \ge 1.37$ (very large effect sizes), assuming 30% dropout, a pre- to posttreatment correlation of 0.18, and a doubling of the SD from pre- to posttreatment. The pre- to posttreatment correlation estimate was based on the mean pooled correlation between pretest and posttest HAM-D scores in 14 CBT trials for adult depression [52], and the estimate of the detec effect size was based on a single degree of freedom contrast in a paired means test implemented in SAS for Windows (version 9.4; SAS Institute).

Feasibility and Acceptability

We examined feasibility and acceptability by reporting (1) dropout rates and reasons (defined as participants not completing an end point HAM-D), (2) patient satisfaction (CSQ), (3) patient feedback (uMARS), (4) patient credibility and expectancy ratings (CEQ), and (5) treatment use. We computed means and SDs for the number of app steps completed, the number of days on which participants completed any actions in the app, the number of messages participants sent their therapist in the app,

the number of sessions participants completed with their therapists, and session time spent per patient per week by therapists. For measures collected at least twice (ie, CSQ, CEQ credibility and expectancy, and treatment use), we used generalized linear mixed models with repeated measures to examine if these self-reported ratings changed over time.

Preliminary Efficacy and Secondary Outcomes

Analyses were first completed using our intent-to-treat sample (participants who completed a baseline assessment) and then repeated with our "per-protocol" sample (participants who completed posttreatment assessments and did not change psychiatric medications or begin psychotherapy during the study; n=24, 86%). We examined the preliminary efficacy of Mindset for Depression plus brief video-conferencing appointments with a therapist on symptoms and well-being outcomes using mixed model analyses with repeated measures (baseline, mid-treatment, posttreatment, and 3-month follow-up) modeled using an unstructured covariance matrix. We then compared pre- to posttreatment differences using a 2-tailed α of .05 to evaluate preliminary efficacy. We similarly compared pretreatment to end of follow-up estimates to estimate whether

changes remained significant by the end of the follow-up. Means are presented as raw means with SDs, while differences between assessments are presented as model-estimated means with CIs (LSM differences [95% CI]) unless otherwise specified. Effect sizes were calculated as Hedges g_{ave} , which takes the correlation of within-participant scores into account [53]. Analyses were conducted for changes in depression symptoms (HAM-D and PHQ-9), functional impairment (WSAS), and quality of life (Q-LES-Q-SF). All analyses were completed using the SAS software (version 9.4, SAS Institute Inc).

Results

Overview

Study participants (N=28) were predominantly female (n=21, 75%), White (n=20, 71%), and single (n=19, 70%), with a mean age of 33.5 (SD 10.9) years. The majority of participants had college or advanced degrees, were employed full time, and came from urban or suburban locations (Tables 2 and 3). Nearly half of the participants (n=13, 45%) had one or more comorbid psychiatric diagnoses, and the average duration of MDD was 15.5 (SD 12.7) years.



 Table 2. Baseline demographics of participants enrolled in the Mindset open trial.

Demographics	Values				
Age (years), mean (SD)	33.5 (10.9)				
Sex at birth, n (%)					
Female	75 (21)				
Male	25 (7)				
Gender identity, n (%)					
Women	75 (21)				
Men	25 (7)				
Sexual orientation, n (%)					
Straight or heterosexual	71 (20)				
Bisexual	18 (5)				
Lesbian, gay, or homosexual	4 (1)				
Other	7 (2)				
Hispanic ethnicity, n (%)	18 (5)				
Race, n (%)					
Asian or Pacific Islander	11 (3)				
Black	4 (1)				
White	71 (20)				
Other	14 (4)				
Education, n (%)					
Less than or equal to a high school graduate	14 (4)				
Technical school or some college	18 (5)				
College graduate	39 (11)				
Graduate or professional school	29 (8)				
Marital status, n (%)					
Single, never married	68 (19)				
Married	18 (5)				
Partnered	7 (2)				
Separated or widowed	7 (2)				
Employment, n (%)					
Full time (≥35 hours per week)	86 (24)				
Student	7 (2)				
Unemployed	4 (1)				
Retired	4 (1)				
Household income (US \$), n (%)					
\$34,999 or less	11 (3)				
\$35,000-74,999	25 (7)				
\$75,000-149,999	50 (14)				
\$150,000 or more	14 (4)				
Geographic location, n (%)					
Urban	43 (12)				
Suburban	46 (13)				
Rural	11 (3)				

https://mental.jmir.org/2024/1/e53998

XSL•FO RenderX

Table 3. Baseline clinical characteristics of participants enrolled in the Mindset open trial.

Characteristics	Values
Duration of MDD ^a (years), mean (SD)	15.5 (12.7)
Current psychiatric comorbidities (DSM-5 ^b diagnoses) ^c , n (%)	
Agoraphobia	7 (2)
Alcohol use disorder	11 (3)
Generalized anxiety disorder	21 (6)
Social anxiety disorder	14 (4)
Other	18 (5)
Number of psychiatric comorbidities, n (%)	
None	54 (15)
1	36 (10)
2	0 (0)
3 or more	11 (3)
Current psychotropic medication ^c , n (%)	
None	57 (16)
SRI ^d	21 (6)
Non-SRI antidepressant	18 (5)
Other psychotropic medication ^e	14 (4)

^aMDD: major depressive disorder.

bDSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

^cPercentage sums may exceed 100% because participants could report more than one diagnosis or be on more than 1 s psychotropic medication.

^dSRI: serotonin reuptake inhibitor.

^eThis included anticonvulsants; no participant reported taking antipsychotics.

Feasibility and Acceptability

Credibility and expectancy scores were moderate to high at preand mid-treatment. The ratings did not differ significantly between timepoints. Mean credibility ratings were 18.9 (SD 3.1) at pretreatment and 19.3 (SD 3.8) at mid-treatment (LSM difference 0.4, 95% CI -1.4 to 2.2; P=.63; g_{ave}=0.13). Mean expectancy ratings were 13.8 (SD 3.3) at pretreatment and 15.0 (SD 5.4) at mid-treatment (LSM difference 1.2, 95% CI -0.7-3.0; P=.22; g_{ave} =0.28). Of the 28 participants, 2 (7%) dropped out of the study prior to their posttreatment assessment at week 8 (Figure 2): 1 during the first week of treatment and 1 after the midpoint assessment; both participants were lost to follow-up despite repeated contact attempts and neither provided a reason for drop out. One more participant was lost to follow-up after the posttreatment assessment. Among the 26 participants who completed the posttreatment assessment, patient satisfaction was high and did not change significantly from mid-treatment (CSQ total score mean 26.3, SD 4.0) to posttreatment (mean 27.2, SD 3.3; LSM difference 1.0, 95% CI -0.1 to 2.2; P=.07; g_{ave} =0.24). Conservatively counting the 2 dropouts as not satisfied, 89% (16/28) were very or mostly (9/28) satisfied and 93% (26/28) would recommend the Mindset for Depression program.

practicing skills from the app on their smartphone and offline for a median of 50 (IQR 30-60) minutes per week up to mid-treatment and 60 (IQR 30-90) minutes per week between mid- and posttreatment. Based on passively collected app use data (n=26), participants accessed the app on 36.8 (SD 10.0) days, completed a median of 7 (IQR 6-8) steps out of 8 steps by the week 8 assessment, and sent a median of 0 (IQR 0-4) between session messages to their therapist through the app. Five participants completed the last assigned step after the week 8 assessment, bringing step completion to a median of 8 (IQR (6-8) by the end of follow-up (3 months), with 58% (15/26) participants completing the final step by then. Participants' overall ratings of the app quality, rated on the 1 (inadequate) to 5 (excellent) scale of the uMARS, was high (mean 4.3, SD 0.4); ratings of the app's functionality (mean 4.5, SD 0.6), aesthetics (mean 4.6, SD 0.4), and information (mean 4.6, SD 0.4) were higher than those of the engagement subscale (mean 3.6, SD 0.6). Participants reported a mean overall star rating of 4.0 (SD (0.5) but were less inclined to endorse that they would be willing to pay for the app (mean 2.5, SD 1.2).

With respect to app use and satisfaction, participants reported

With respect to therapist support, participants attended an average of 7.6 (SD 1.5) of the possible 8 brief sessions, each of which lasted approximately 24.5 (SD 1.1) minutes. In the sessions, therapists mainly covered behavioral strategies (mean 10.9, SD 3.4 minutes), cognitive strategies (mean 6.9, SD 2.7

minutes), psychoeducation (mean 2.5, SD 2.9 minutes), and mindfulness strategies (mean 2.4, SD 1.2 minutes), with only a little time (<2 minutes on average) spent on explicit

motivational strategies, risk management, and technical issues. Participants had a mean homework completion rate (per therapist report) of 82.7% (SD 13.8%).

Figure 2. Flow of participants through the 8-week open trial of the Mindset Depression app with brief therapist visits on the web for people with a primary diagnosis of major depressive disorder. Reasons for ineligibility include diagnosis of bipolar disorder or severe substance use disorder, PHQ-9 score < 10, past CBT for MDD, acute, active suicidal ideation, and MDD not being the primary diagnosis. CBT: cognitive behavioral therapy.



Preliminary Efficacy and Secondary Outcomes

Over the course of the 8-week treatment, participants' depression severity decreased significantly on both the clinician-rated (HAM-D: P<.001; $g_{ave}=1.47$) and self-reported measures (PHQ-9: P<.001; $g_{ave}=1.89$; 4). Concurrently, participants' self-rated functional impairment decreased (WSAS: P<.001; $g_{ave}=1.29$), and their self-rated quality of life increased (Q-LES-Q-SF: P<.001; $g_{ave}=1.74$). These changes persisted through the 3-month follow-up, with effect sizes remaining largely the same (Table 4). The results did not differ meaningfully in the per-protocol analyses; HAM-D, PHQ-9, WSAS, and Q-LES-Q scores all improved with statistically significant and large effect sizes (Table 5).

Conservatively counting dropouts as having not responded to treatment, 46% (13/28) participants responded and another 7%

(2/28) partially responded at posttreatment. Regarding remission, and again counting dropouts also as not remitting, 36% (10/28) achieved remission at posttreatment. Based on the per-protocol sample, 54% (13/24) participants fully responded to treatment and 8% (2/24) partially responded to treatment; 42% (10/24) achieved remission by the posttreatment assessment. By the end of follow-up and counting dropouts as not responding and not remitting, 50% (14/28) participants had responded to treatment, and an additional 21% (6/28) of participants had partially responded to treatment; 36% (10/28) participants were in remission. In the per-protocol sample at the end of follow-up and excluding the additional participant lost to follow-up, 61% (14/23) of participants had responded to treatment, and 26% (6/23) of participants had partially responded to treatment; 43% (10/23) participants were in remission at the 3-month follow-up assessment.



Wilhelm et al

Table 4. Baseline, mid-treatment (week 4), end of treatment (week 8), and follow-up (week 20) estimated mean scores on key clinical outcome measures.

Outcome measure	Baseline, LSM ^a (SE)	Week 8, LSM (SE)	Week 20, LSM (SE)	Estimated differ- ence (week 0-8), LSM (95% CI)	P value	Effect size (week 0-8) ^b , Hedges g _{av}	Estimated differ- ence (week 0-20), LSM (95%CI)	P value	Effect size (week 0-20) ^a , Hedges g_{av}
HAM-D ^c total scores	19.1 (0.9)	11.3 (1.2)	10.4 (1.6)	-7.8 (-10.5 to -5.2)	<.001	1.47	-8.7 (-12.0 to -5.4)	<.001	1.44
PHQ-9- ^d total scores	15.1 (0.7)	7.1 (1.0)	6.9 (0.9)	-8.0 (-9.9 to -6.1)	<.001	1.89	-8.2 (-10.2 to -6.3)	<.001	1.97
WSAS ^e total scores	23.1 (1.5)	13.2 (1.5)	12.5 (1.8)	-9.9 (-13.5 to -6.4)	<.001	1.29	-10.6 (-14.5 to -6.8)	<.001	1.26
Q-LES-Q-SF ^f % scores	40.9 (2.2)	62.5 (2.6)	62.8 (2.8)	21.5 (14.5 to 28.6)	<.001	1.74	21.9 (15.0 to 28.8)	<.001	1.73

^aLSM: least squares mean.

^bWithin-group effect sizes were calculated as Hedges g_{ave} for differences from baseline to week 8 or 20, respectively, using raw means data.

^cHAM-D: Hamilton Depression Rating Scale (score range 0 to 52, where higher scores indicate greater depression severity).

^dPHQ-9: Patient Health Questionnaire-9 item (score range 0 to 27, where higher scores indicate greater depression severity).

^eWSAS: Work and Social Adjustment Scale (score range 0 to 40, where higher scores mean higher functional impairment).

^fQ-LES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire- Short Form (percent score range 0% to 100%, where higher scores correspond to greater ratings of quality of life).

Table 5. Baseline, mid-treatment (week 4), end-of-treatment (week 8), and follow-up (week 20) estimated mean scores on key clinical outcome measures in the per-protocol sample (n=24).

Outcome measure	Baseline, LSM ^a (SE)	Week 8, LSM (SE)	Week 20, LSM (SE)	Estimated differ- ence (week 0- 8), LSM (95% CI)	P value	Effect size (week 0- 8) ^b , Hedges g _{av}	Estimated differ- ence (week 0-20), LSM (95% CI)	P value	Effect size (week 0- 20) ^a , Hedges <i>gav</i>
HAM-D ^c total scores	18.7 (1.0)	10.3 (1.2)	8.6 (1.3)	-8.4 (-11.1 to -5.6)	<.001	1.52	-10.1 (-13.1 to -7.1)	<.001	1.79
PHQ-9 ^d total scores	14.7 (0.8)	6.4 (1.0)	6.1 (0.9)	-8.3 (-10.3 to -6.2)	<.001	1.89	-8.6 (-10.6 to -6.6)	<.001	2.03
WSAS ^e total scores	22.0 (1.6)	12.1 (1.4)	10.7 (1.6)	-9.8 (-13.7 to -6.0)	<.001	1.32	-11.2 (-15.2 to -7.3)	<.001	1.42
Q-LES-Q-SF ^f % scores	41.7 (2.3)	63.7 (2.8)	65.3 (2.7)	22.0 (14.5 to 29.6)	<.001	1.73	23.6 (16.7 to 30.5)	<.001	1.92

^aLSM = least squares mean.

^bWithin-group effect sizes were calculated as Hedges g_{ave} for differences from baseline to week 8 or 20, respectively, using raw means for completers only.

^cHAM-D: Hamilton Depression Rating Scale (score range 0 to 52, where higher scores indicate greater depression severity).

^dPHQ-9: Patient Health Questionnaire-9 item (score range 0 to 27, where higher scores indicate greater depression severity).

^eWSAS: Work and Social Adjustment Scale (score range 0 to 40, where higher scores mean higher functional impairment).

^fQ-LES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form (percent score range 0% to 100%, where higher scores correspond to greater ratings of quality of life).

Adverse Events and Medication Changes

Overall, 17 out of 28 participants reported a total of 33 adverse events during the 8-week treatment phase of the trial in categories such as psychiatric symptoms (n=21, 64%; eg, increased suicidal ideation, depression, or anxiety, sleep difficulties, and emotional distress), infections (n=6, 18%; eg, COVID-19, shingles, and illness), physical injuries (n=3, 9%), and other (n=3, 9%). All adverse events were identified as either mild (new event that did not interfere with activities of daily living; 25/33, 75.8%) or moderate (new event that posed some interference or required intervention to prevent interference; 8/33, 24.2%). No serious adverse events occurred in this trial. By assessing how likely it was that reported adverse events were related to treatment, most events were found to be definitely unrelated (19/33, 57.6%), followed by unlikely to be related (5/33, 15.2%) or possibly related (9/33, 27.3%). A waxing and waning course of MDD symptoms and suicidal ideation is common in MDD. However, there were no adverse events indicating significant clinical deterioration in the trial

and no principal investigator–initiated withdrawals. Two participants changed psychotropic medications during the treatment phase: 1 participant increased the dosage of their medication for anxiety and 1 participant discontinued their medication for depression. There were no reported therapy changes during the treatment phase. In the 3-month follow-up phase of the trial, 7 participants reported an additional 10 adverse events in the categories of psychiatric symptoms (4/10, 40%), general disorders (2/10, 20%), and other (4/10, 40%), which were found to be definitely unrelated (5/10, 50%), unlikely to be related (1/10, 10%), and possibly related (4/10, 40%) to treatment. Also during the follow-up period, 5 participants changed psychotropic medications and 3 participants started individual therapy or counseling (non-CBT; for mood or anxiety, and traumatic event or PTSD).

Discussion

Principal Findings

In this study, we examined the feasibility, acceptability, and preliminary clinical impact of Mindset for Depression, an 8-week app-based CBT with therapist support. The results support Mindset for Depression as a viable treatment option for individuals with moderate to severe MDD. Treatment was feasible to deliver in a setting and acceptable to patients who varied widely in age, severity of the symptoms, and other clinical and demographic dimensions, as indicated by high retention rates (27/29, 93%), favorable satisfaction ratings (CSQ), and positive user feedback (uMARS). The results also showed that the treatment was efficacious. There was a significant reduction in clinician-rated (HAM-D) depression severity with a large effect size as well as significant improvement in functioning and quality of life. After just 8 weeks, about half of the participants were rated as treatment responders and a third were in remission, and these changes were maintained throughout the 3-month follow-up. These results are similar to face-to-face psychotherapy [54] and comparable to guided internet-delivered CBTs, notable given the short treatment duration, younger age, and relatively higher severity of the sample, all of which are associated with lower odds of response and remission in digital treatment [55].

Encouragingly, app ratings were above average for mental health apps [56]. Compared to other mHealth for depression or anxiety [57-60] and treatment-as-usual for depression [61], overall treatment satisfaction scores (CSQ) were also excellent. This was achieved despite clinician time that was below typical courses of CBT for depression (8 sessions averaging 24 minutes vs upwards of 20 sessions lasting on average 45-50 minutes) [62]. Indeed, much of the time-consuming didactic content (eg, psychoeducation about the CBT model) was administered by the app through readings, videos, and practice questions, conserving clinician time for more personalized skills' review, practice, and tailoring and addressing risk issues. App usage data were excellent, with most participants reaching the final step by the end of 8 weeks as intended and reporting regularly practicing skills on or off the app each week. Moreover, participants rarely used the messaging function between sessions (eg, seeking additional clarification or encouragement), which

would be unbillable clinician time. In this way, Mindset for Depression has the potential to improve the reach of CBT therapists and hopefully reduce treatment gaps, particularly for underserved communities [63,64].

These results are important because the high cost and limited availability of trained clinicians are major barriers to the dissemination of traditional psychotherapy [65]. Our findings add to an emerging literature demonstrating the potential of guided smartphone-based CBT to mitigate these challenges. However, although numerous seemingly efficacious therapist-supported digital treatments have been created for depression [10], few are available outside of research settings or integrated into a health care system [25]. Created in collaboration with an industry partner to accelerate the dissemination pipeline and allow for ongoing technical maintenance and improvements, Mindset for Depression is commercially available and poised to be truly scalable and successful in real clinical settings. Setting it apart, Mindset for Depression was collaboratively developed with a design team, clinicians, and people with lived experience as well as rigorously applied user interface and experience best practices for mobile platforms. Critically, "users" in the user-centered design process included both patients and clinicians. This approach aligns clinical and engagement incentives so that one is not delivered at the expense of the other and yields an easy to use, streamlined, and effective treatment program. Concretely, this translated to pacing content (delivering or unlocking intervention components in a stepwise manner to encourage practice and mastery of concepts and skills that build on one another), shorter activity lengths, creating a professional and approachable tone, and inclusion of feedback loops. As standardized content is delivered via the app in each step, supporting clinicians are able to prioritize personalizing treatment and use their specialized skillsets, such as addressing unique barriers to motivation, engagement, or response. Critical next steps would be to directly evaluate the program and its readiness to scale in larger scale effectiveness trials and real-world settings.

The results also provide important guidance for improving the program. First, participant feedback (eg, uMARS engagement subscale) indicates that increased customization and interactivity could improve the app's appeal. This is consistent with the larger literature showing user preferences for apps with such features and negative reactions to apps whose content is repetitive and not personally relevant [66]. Although most participants shared positive views of the Mindset app, including indicating that they would recommend the app to friends, there was a mixed response regarding their willingness to pay for the Mindset app. It is unclear to what extent this reflects (1) that this question was asked after treatment and thus patients no longer felt the need to use the app; patients were meant to complete all therapeutic content within the 8-week treatment period; (2) a gap between what patients find beneficial and what they are willing to pay for; other studies have similarly found a reluctance to pay for mental health apps [66,67]; or (3) whether the app and concurrent therapist support were experienced as critically linked, and thus the app alone was not as valued. Indeed, half of participants indicated weekly brief sessions were the exact right amount of therapist contact and only 2 would

XSL•FO RenderX

have preferred less contact. Resolving this question will be necessary for developing a commercially sustainable implementation plan. Moreover, future iterations would benefit from broadening outcomes of interest. For example, beyond reducing symptoms of depression or other mental health concerns, an optimal intervention would also foster positive emotions and thriving and perhaps target common comorbidities, such as sleep difficulties or substance use. These outcomes should be captured in future studies and additional treatment components integrated as appropriate.

Limitations

The study has limitations that should be considered. First, the study has the inherent limitation of an open trial. Without a control group, we cannot conclusively determine that the treatment causes improvements in symptoms. Future controlled large-scale trials are needed. Second, the patient sample was self-selected, recruitment platforms were diverse, and our study therapists were trained in the use of digital therapeutics. Thus, patient and clinician stakeholders might have been biased toward individuals who are motivated by app-based therapy. Future research in real-world clinical settings is warranted. Third, although the large proportion of White women in the sample is consistent with past work and higher MDD prevalence and rate of treatment seeking in women [10], greater representation of

patients with other racial and gender identities would strengthen our conclusions and ongoing treatment improvements. Fourth, we had adequate power to detect moderate to large treatment effects; a larger replication is needed to explore moderators and mediators, which are important for tiered care models. Finally, a longer follow-up period and health economics metrics would be required to see the full time and cost-savings potential of Mindset for Depression.

Conclusions

Mindset for Depression offers flexible app-led psychoeducation, skills practice, and support to patients with complementary clinician guidance to promote sustained engagement, monitor safety, and tailor treatment further to individual patient needs. The findings show that Mindset for Depression is a feasible, acceptable, and efficacious tool for adults with MDD. The hope is that such a program could be one cost-effective solution to barriers to psychotherapy dissemination and significantly increase access to evidence-based care. Although these initial results are very promising, more work remains to personalize the amount of therapist support and dose of treatment individuals receive to optimize treatment and increase rates of response and remission. The next steps include testing Mindset for Depression in a fully powered randomized controlled trial as well as the real-world clinical settings in which it is deployed.

Acknowledgments

The research was supported by Koa Health. Investigators from Massachusetts General Hospital (MGH) (SW, EEB, KHB, JLG, and HW) developed the Mindset app in collaboration with technologists and designers from Koa Health via a user-centered, iterative design approach. The investigators from MGH were responsible for the study design and the execution of the study. Koa Health supported the technical activities related to the development and deployment of the app. Koa Health had no role in the recruitment of participants and had no access to data during the trial. The sponsor was able to make minor edits to the draft of the paper. HW was supported in part by the National Institute of Mental Health of the National Institutes of Health under award number K23MH119372. KHB was supported in part by the National Institute of Mental Health of the National Institutes of Health under award number K23MH120436. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank the following individuals for their invaluable work on the project: Lauren Fisher, Margaret Hall, Brynn Huguenel, Adam Jaroszewski, Geneva Jonathan, Jennifer Ragan, and Jasmine Williams.

Conflicts of Interest

SW has received royalties from Elsevier Publications, Guilford Publications, New Harbinger Publications, Springer, and Oxford University Press. SW has also received speaking honoraria from various academic institutions and foundations, including a research award from the National Alliance on Mental Illness. In addition, she received honoraria for her role on the Scientific Advisory Board for One-Mind (PsyberGuide), Koa Health Inc, Noom Inc, and Jimini Health. SW has received research support from Koa Health. EEB has received research support from Koa Health. EEB has a consulting agreement with Otsuka Pharmaceutical Development & Commercialization Inc and is on the Scientific Advisory Board for AugMend Health Inc. KHB has received research support from Koa Health. IS has received research support from Koa Health. SSH has received research support from Koa Health. DK has received research support from Koa Health. JLG has received research support from Koa Health. She has received speaking honoraria from L'Oreal (for a presentation at a SkinCeuticals cosmetic surgery and dermatology conference) and RBC Consultants for the CeraVe Psychodermatology Advisory Board. HW has received research support from Koa Health and the National Institute of Mental Health (NIMH; K23MH119372). HW has a consulting agreement with Hello Therapeutics, Inc. THM has received research support from Koa Health. He has received unrelated editorial honoraria from Springer Nature as well as grants to his institution from InterSystems, National Institute of Mental Health, National Institute of Nursing Research, National Institute on Aging, and National Human Genome Research Institute. OH is the founder/chief executive officer of Koa Health, a digital mental health company that collaborated with SW and her team at MGH to build Mindset. OH also serves on the World Health Organization Roster of Experts for Digital Health, sits on the Board of EMPOWER (a nonprofit organization promoting the training of community health workers to provide mental health care), and is a member of the expert panel for

implementing the Wellcome Trust's mental health strategy. Dr Harrison is a Royal Society Entrepreneur in Residence in healthcare Artificial Intelligence (AI) at Oxford University.

Multimedia Appendix 1

Mindset Therapist Adherence Scale. [DOCX File , 20 KB-Multimedia Appendix 1]

References

- 1. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. Arlington, VA. American Psychiatric Publishing; 2013.
- 2. 2020 National Survey on Drug Use and Health (NSDUH) releases. Substance Abuse and Mental Health Services Administration. 2020. URL: <u>https://www.samhsa.gov/data/release/2020-national-survey-drug-use-and-health-nsduh-releases</u> [accessed 2024-03-12]
- Ettman CK, Cohen GH, Abdalla SM, Sampson L, Trinquart L, Castrucci BC, et al. Persistent depressive symptoms during COVID-19: a national, population-representative, longitudinal study of U.S. adults. Lancet Reg Health Am. 2022;5:100091.
 [FREE Full text] [doi: 10.1016/j.lana.2021.100091] [Medline: 34635882]
- 4. Friedrich MJ. Depression is the leading cause of disability around the world. JAMA. 2017;317(15):1517. [doi: 10.1001/jama.2017.3826] [Medline: 28418490]
- 5. Greenberg PE, Fournier AA, Sisitsky T, Simes M, Berman R, Koenigsberg SH, et al. The economic burden of adults with major depressive disorder in the United States (2010 and 2018). Pharmacoeconomics. 2021;39(6):653-665. [FREE Full text] [doi: 10.1007/s40273-021-01019-4] [Medline: 33950419]
- Moitra M, Santomauro D, Collins PY, Vos T, Whiteford H, Saxena S, et al. The global gap in treatment coverage for major depressive disorder in 84 countries from 2000-2019: a systematic review and Bayesian meta-regression analysis. PLoS Med. 2022;19(2):e1003901. [FREE Full text] [doi: 10.1371/journal.pmed.1003901] [Medline: 35167593]
- López-López JA, Davies SR, Caldwell DM, Churchill R, Peters TJ, Tallon D, et al. The process and delivery of CBT for depression in adults: a systematic review and network meta-analysis. Psychol Med. 2019;49(12):1937-1947. [FREE Full text] [doi: 10.1017/S003329171900120X] [Medline: 31179960]
- Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychol Med. 2014;44(6):1303-1317. [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
- Webb CA, Rosso IM, Rauch SL. Internet-based cognitive-behavioral therapy for depression: current progress and future directions. Harv Rev Psychiatry. 2017;25(3):114-122. [FREE Full text] [doi: 10.1097/HRP.00000000000139] [Medline: 28475503]
- 10. Børtveit L, Dechsling A, Sütterlin S, Nordgreen T, Nordahl-Hansen A. Guided internet-delivered treatment for depression: scoping review. JMIR Ment Health. 2022;9(10):e37342. [FREE Full text] [doi: 10.2196/37342] [Medline: 36194467]
- 11. Torous J, Levin ME, Ahern DK, Oser ML. Cognitive behavioral mobile applications: clinical studies, marketplace overview, and research agenda. Cogn Behav Pract. 2017;24(2):215-225. [doi: 10.1016/j.cbpra.2016.05.007]
- 12. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry. 2019;18(3):325-336. [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
- Mohr DC, Azocar F, Bertagnolli A, Choudhury T, Chrisp P, Frank R, et al. Banbury forum consensus statement on the path forward for digital mental health treatment. Psychiatr Serv. 2021;72(6):677-683. [FREE Full text] [doi: 10.1176/appi.ps.202000561] [Medline: <u>33467872</u>]
- Wasil AR, Gillespie S, Shingleton R, Wilks CR, Weisz JR. Examining the reach of smartphone apps for depression and anxiety. Am J Psychiatry. 2020;177(5):464-465. [FREE Full text] [doi: 10.1176/appi.ajp.2019.19090905] [Medline: 32354266]
- Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of Intellicare: a publicly available suite of mental health and well-being apps. Internet Interv. 2016;4(2):152-158. [FREE Full text] [doi: 10.1016/j.invent.2016.06.003] [Medline: 27398319]
- Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res. 2019;21(9):e14567. [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
- 17. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res. 2011;13(1):e30. [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
- 18. Mohr DC, Schueller SM, Tomasino KN, Kaiser SM, Alam N, Karr C, et al. Comparison of the effects of coaching and receipt of app recommendations on depression, anxiety, and engagement in the IntelliCare platform: factorial randomized controlled trial. J Med Internet Res. 2019;21(8):e13609. [FREE Full text] [doi: 10.2196/13609] [Medline: 31464192]

- Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health. 2018;21(3):116-119. [FREE Full text] [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
- 20. Muñoz RF. The efficiency model of support and the creation of digital apothecaries. Clin Psychol (New York). 2017;24(1):46-49. [doi: 10.1111/cpsp.12174]
- 21. Andersson G, Topooco N, Havik O, Nordgreen T. Internet-supported versus face-to-face cognitive behavior therapy for depression. Expert Rev Neurother. 2016;16(1):55-60. [doi: <u>10.1586/14737175.2015.1125783</u>] [Medline: <u>26610160</u>]
- 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;13(3):288-295.
 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
- 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;47(1):1-18. [FREE Full text] [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
- 24. Shechter A, Diaz F, Moise N, Anstey DE, Ye S, Agarwal S, et al. Psychological distress, coping behaviors, and preferences for support among New York healthcare workers during the COVID-19 pandemic. Gen Hosp Psychiatry. 2020;66:1-8. [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.06.007] [Medline: 32590254]
- Bernstein EE, Weingarden H, Wolfe EC, Hall MD, Snorrason I, Wilhelm S. Human support in app-based cognitive behavioral therapies for emotional disorders: scoping review. J Med Internet Res. 2022;24(4):e33307. [FREE Full text] [doi: 10.2196/33307] [Medline: 35394434]
- 26. Chang S, Gray L, Torous J. Smartphone app engagement and clinical outcomes in a hybrid clinic. Psychiatry Res. 2023;319:115015. [doi: 10.1016/j.psychres.2022.115015] [Medline: 36549096]
- 27. Lakhtakia T, Torous J. Current directions in digital interventions for mood and anxiety disorders. Curr Opin Psychiatry. 2022;35(2):130-135. [doi: 10.1097/YCO.000000000000772] [Medline: 34966117]
- Karyotaki E, Efthimiou O, Miguel C, Bermpohl FMG, Furukawa TA, Cuijpers P, Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration, et al. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. JAMA Psychiatry. 2021;78(4):361-371. [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
- 29. Depression in adults: treatment and management [NICE guideline NG222]. National Institute for Health and Care Excellence. URL: <u>https://www.nice.org.uk/guidance/ng222/</u> [accessed 2024-03-12]
- 30. Posner K, Brent D, Lucas C, Gould M, Stanley B, Brown G, et al. Columbia-Suicide Severity Rating Scale (C-SSRS). The Columbia Lighthouse Project. 2008. URL: <u>https://cssrs.columbia.edu/wp-content/uploads/C-SSRS_Pediatric-SLC_11.14.</u> <u>16.pdf</u> [accessed 2024-03-12]
- 31. Forde F, Frame M, Hanlon P, MacLean G, Nolan D, Shajahan P, et al. Optimum number of sessions for depression and anxiety. Nurs Times. 2005;101(43):36-40. [Medline: <u>16276843</u>]
- 32. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. Psychiatr Serv. 2017;68(5):427-429. [FREE Full text] [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
- 33. Ciharova M, Furukawa TA, Efthimiou O, Karyotaki E, Miguel C, Noma H, et al. Cognitive restructuring, behavioral activation and cognitive-behavioral therapy in the treatment of adult depression: a network meta-analysis. J Consult Clin Psychol. 2021;89(6):563-574. [doi: 10.1037/ccp0000654] [Medline: 34264703]
- 34. Sauer-Zavala S, Bentley KH, Steele SJ, Tirpak JW, Ametaj AA, Nauphal M, et al. Treating depressive disorders with the Unified Protocol: a preliminary randomized evaluation. J Affect Disord. 2020;264:438-445. [FREE Full text] [doi: 10.1016/j.jad.2019.11.072] [Medline: <u>31759672</u>]
- 35. Beck AT. Cognitive Therapy of Depression. New York. Guilford Publications; 1979.
- Bentley KH, Bernstein EE, Wallace B, Mischoulon D. Treatment for anxiety and comorbid depressive disorders: transdiagnostic cognitive-behavioral strategies. Psychiatr Ann. 2021;51(5):226-230. [FREE Full text] [doi: 10.3928/00485713-20210414-01] [Medline: <u>34433988</u>]
- Dimidjian S, Martell CR, Herman-Dunn R, Hubley S. Behavioral activation for depression. In: Barlow DH, editor. Clinical Handbook of Psychological Disorders: A Step-by-Step Treatment Manual, 5th Edition. New York. Guilford Publications; 2014;353-393.
- 38. Sturmey P. Behavioral activation is an evidence-based treatment for depression. Behav Modif. 2009;33(6):818-829. [doi: 10.1177/0145445509350094] [Medline: 19933444]
- MacKenzie MB, Kocovski NL. Mindfulness-based cognitive therapy for depression: trends and developments. Psychol Res Behav Manag. 2016;9:125-132. [FREE Full text] [doi: 10.2147/PRBM.S63949] [Medline: 27274325]
- 40. Akçayır G, Akçayır M. The flipped classroom: a review of its advantages and challenges. Comput Educ. 2018;126:334-345. [doi: <u>10.1016/j.compedu.2018.07.021</u>]
- 41. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998;59(Suppl 20):22-33;quiz 34. [Medline: <u>9881538</u>]

- 42. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry. 2000;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
- Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth. 2016;4(2):e72. [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
- 44. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann. 1982;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]
- 45. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960;23(1):56-62. [FREE Full text] [doi: 10.1136/jnnp.23.1.56] [Medline: 14399272]
- 46. Rosenblat JD, Lee Y, McIntyre RS. The effect of pharmacogenomic testing on response and remission rates in the acute treatment of major depressive disorder: a meta-analysis. J Affect Disord. 2018;241:484-491. [doi: <u>10.1016/j.jad.2018.08.056</u>] [Medline: <u>30149336</u>]
- 47. Nierenberg AA, DeCecco LM. Definitions of antidepressant treatment response, remission, nonresponse, partial response, and other relevant outcomes: a focus on treatment-resistant depression. J Clin Psychiatry. 2001;62(Suppl 16):5-9. [Medline: 11480882]
- 48. Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. Arch Gen Psychiatry. 1991;48(9):851-855. [doi: 10.1001/archpsyc.1991.01810330075011] [Medline: 1929776]
- 49. Mundt JC, Marks IM, Shear MK, Greist JH. The work and social adjustment scale: a simple measure of impairment in functioning. Br J Psychiatry. 2002;180:461-464. [FREE Full text] [doi: 10.1192/bjp.180.5.461] [Medline: 11983645]
- 50. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of life enjoyment and satisfaction questionnaire: a new measure. Psychopharmacol Bull. 1993;29(2):321-326. [Medline: <u>8290681</u>]
- 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;282(18):1737-1744. [FREE Full text] [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
- 52. Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJH. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry. 2016;15(3):245-258. [FREE Full text] [doi: 10.1002/wps.20346] [Medline: 27717254]
- 53. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;4:863. [FREE Full text] [doi: 10.3389/fpsyg.2013.00863] [Medline: 24324449]
- Cuijpers P, Karyotaki E, Weitz E, Andersson G, Hollon SD, van Straten A. The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. J Affect Disord. 2014;159:118-126. [doi: 10.1016/j.jad.2014.02.026] [Medline: 24679399]
- 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;63:80-92. [doi: <u>10.1016/j.cpr.2018.06.007</u>] [Medline: <u>29940401</u>]
- 56. Hudson G, Negbenose E, Neary M, Jansli SM, Schueller SM, Wykes T, et al. Comparing professional and consumer ratings of mental health apps: mixed methods study. JMIR Form Res. 2022;6(9):e39813. [FREE Full text] [doi: 10.2196/39813] [Medline: 36149733]
- Schlicker S, Baumeister H, Buntrock C, Sander L, Paganini S, Lin J, et al. 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. [FREE Full text] [doi: 10.2196/16398] [Medline: 32293577]
- Netter AL, Beintner I, Brakemeier EL. Adding an app-based intervention to the cognitive behavioral analysis system of psychotherapy in routine outpatient psychotherapy treatment: proof-of-concept study. JMIR Form Res. 2022;6(8):e35482.
 [FREE Full text] [doi: 10.2196/35482] [Medline: 35943764]
- 59. Sharma G, Schlosser L, Jones BDM, Blumberger DM, Gratzer D, Husain MO, et al. Brief app-based cognitive behavioral therapy for anxiety symptoms in psychiatric inpatients: feasibility randomized controlled trial. JMIR Form Res. 2022;6(11):e38460. [FREE Full text] [doi: 10.2196/38460] [Medline: 36322113]
- 60. Gotthardt M, Striegl J, Loitsch C, Weber G. Voice assistant-based CBT for depression in students: effects of empathy-driven dialog management. In: Miesenberger K, Kouroupetroglou G, Mavrou K, Manduchi R, Rodriguez MC, Penáz P, editors. Computers Helping People with Special Needs. ICCHP-AAATE 2022. Lecture Notes in Computer Science, Vol 13341. Cham. Springer; 2022;451-461.
- 61. Richards DA, Hill JJ, Gask L, Lovell K, Chew-Graham C, Bower P, et al. Clinical effectiveness of collaborative care for depression in UK primary care (CADET): cluster randomised controlled trial. BMJ. 2013;347:f4913. [FREE Full text] [doi: 10.1136/bmj.f4913] [Medline: 23959152]

- 62. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Can J Psychiatry. 2013;58(7):376-385. [FREE Full text] [doi: 10.1177/070674371305800702] [Medline: 23870719]
- 63. Ramos G, Chavira DA. Use of technology to provide mental health care for racial and ethnic minorities: evidence, promise, and challenges. Cogn Behav Pract. 2022;29(1):15-40. [doi: 10.1016/j.cbpra.2019.10.004]
- Bae H, Shin H, Ji HG, Kwon JS, Kim H, Hur JW. App-based interventions for moderate to severe depression: a systematic review and meta-analysis. JAMA Netw Open. 2023;6(11):e2344120. [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.44120] [Medline: 37983028]
- 65. Cavanagh K. Geographic inequity in the availability of cognitive behavioural therapy in England and Wales: a 10-year update. Behav Cogn Psychother. 2014;42(4):497-501. [doi: 10.1017/S1352465813000568] [Medline: 23829940]
- Balaskas A, Schueller SM, Cox AL, Rashleigh C, Doherty G. Examining young adults daily perspectives on usage of anxiety apps: a user study. PLOS Digit Health. 2023;2(1):e0000185. [FREE Full text] [doi: 10.1371/journal.pdig.0000185] [Medline: 36812622]
- 67. Bakker D, Kazantzis N, Rickwood D, Rickard N. Development and pilot evaluation of smartphone-delivered cognitive behavior therapy strategies for mood- and anxiety-related problems: MoodMission. Cogn Behav Pract. 2018;25(4):496-514. [doi: 10.1016/j.cbpra.2018.07.002]

Abbreviations

CBT: cognitive behavioral therapy CEQ: Credibility/Expectancy Questionnaire CSQ: Client Satisfaction Questionnaire DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition HAM-D: Hamilton Depression Rating Scale HIPAA: Health Insurance Portability and Accountability Act MDD: major depressive disorder PHQ-9: Patient Health Questionnaire-9 Q-LES-Q-SF: The Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form SMART: specific, measurable, achievable, relevant, and time-bound uMARS: Mobile Application Rating Scale User Version WSAS: Work and Social Adjustment Scale

Edited by J Torous; submitted 26.10.23; peer-reviewed by A Brahmer, NH Zainal; comments to author 25.11.23; revised version received 19.12.23; accepted 03.01.24; published 09.04.24

Please cite as:

Wilhelm S, Bernstein EE, Bentley KH, Snorrason I, Hoeppner SS, Klare D, Greenberg JL, Weingarden H, McCoy TH, Harrison O Feasibility, Acceptability, and Preliminary Efficacy of a Smartphone App–Led Cognitive Behavioral Therapy for Depression Under Therapist Supervision: Open Trial JMIR Ment Health 2024;11:e53998 URL: https://mental.jmir.org/2024/1/e53998

doi: <u>10.2196/53998</u> PMID: <u>38592771</u>

©Sabine Wilhelm, Emily E Bernstein, Kate H Bentley, Ivar Snorrason, Susanne S Hoeppner, Dalton Klare, Jennifer L Greenberg, Hilary Weingarden, Thomas H McCoy, Oliver Harrison. Originally published in JMIR Mental Health (https://mental.jmir.org), 09.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on https://mental.jmir.org/, as well as this copyright and license information must be included.

