

RESEARCH ARTICLE

ANXIETY AND DEPRESSION
ASSOCIATION OF AMERICA

WILEY

Efficacy of digital cognitive behavioral therapy for moderate-to-severe symptoms of generalized anxiety disorder: A randomized controlled trial

Jenna R. Carl^{1,2} | Christopher B. Miller^{1,2,3}  | Alasdair L. Henry^{1,2,3} |
Michelle L. Davis^{1,2} | Richard Stott^{1,2,4} | Jasper A. J. Smits⁵ | Richard Emsley⁶ |
Jenny Gu^{1,2,4} | Olivia Shin^{1,2} | Michael W. Otto⁷ | Michelle G. Craske⁸  |
Kate E. A. Saunders^{4,9} | Guy M. Goodwin⁴ | Colin A. Espie^{1,2,3}

¹Big Health Inc., San Francisco, California²Big Health Inc., London, UK³Nuffield Department of Clinical
Neurosciences, Sleep and Circadian
Neuroscience Institute, University of Oxford,
Oxford, UK⁴Department of Psychiatry, University of
Oxford, Oxford, UK⁵Department of Psychology, The University of
Texas at Austin, Austin, Texas⁶Department of Biostatistics and Health
Informatics, Institute of Psychiatry,
Psychology, & Neuroscience, King's College
London, London, UK⁷Department of Psychological and Brain
Sciences, Boston University, Boston,
Massachusetts⁸Anxiety and Depression Research Centre
(ADRC), University of California,
Los Angeles, California⁹Oxford Health NHS Foundation Trust,
Warneford Hospital, Oxford, UK**Correspondence**

Christopher B. Miller, Big Health, Runway East
London Bridge, 20 St. Thomas St.,
London SE1 9RS, UK.
Email: chris.miller@bighealth.com

Abstract

Background: Cognitive behavioral therapy (CBT) is an efficacious intervention for generalized anxiety disorder (GAD). Digital CBT may provide a scalable means of delivering CBT at a population level. We investigated the efficacy of a novel digital CBT program in those with GAD for outcomes of anxiety, worry, depressive symptoms, sleep difficulty, wellbeing, and participant-specific quality of life.

Methods: This online, two-arm parallel-group superiority randomized controlled trial compared digital CBT with waitlist control in 256 participants with moderate-to-severe symptoms of GAD. Digital CBT (*Daylight*), was delivered using participants' own smartphones. Online assessments took place at baseline (Week 0; immediately preceding randomization), mid-intervention (Week 3; from randomization), post-intervention (Week 6; primary endpoint), and follow-up (Week 10).

Results: Overall, 256 participants were randomized and intention-to-treat analysis found *Daylight* reduced symptoms of anxiety compared with waitlist control at post-intervention, reflecting a large effect size (adjusted difference [95% CI]: 3.22 [2.14, 4.31], $d = 1.08$). Significant improvements were found for measures of worry; depressive symptoms, sleep difficulty, wellbeing, and participant-specific quality of life.

Conclusion: Digital CBT (*Daylight*) appears to be safe and efficacious for symptoms of anxiety, worry, and further measures of mental health compared with waitlist control in individuals with GAD.

KEYWORDS

anxiety, cognitive behavioral therapy, generalized anxiety disorder, internet, web-based

1 | INTRODUCTION

Generalized anxiety disorder (GAD) is characterized by excessive anxiety and worry about a variety of events or activities that is difficult to control (American Psychiatric Association, 2013). GAD affects ~4% of the population (Kessler, Petukhova, Sampson, Zaslavsky, & Wittchen, 2012). It impacts wellbeing, life satisfaction, and health status; and leads to increased healthcare utilization and decreased work productivity (Revicki et al., 2012). Pharmacotherapy and cognitive behavioral therapy (CBT) are considered first-line interventions for GAD (Anxiety and Depression Association of America, 2020; Canadian Psychiatric Association, 2006; Locke, Kirst, & Shultz, 2015; National Institute for Health and Clinical Excellence, 2011). Meta-analyses show CBT has large intervention effects relative to control conditions (Carpenter et al., 2018) and is more efficacious than pharmacotherapy (Carl et al., 2020).

Barriers limit access to CBT including insufficient numbers of trained therapists, intervention costs, waiting lists, incompatible scheduling, distance from services, and perceived stigma (Comer & Barlow, 2014; Gunter & Whittal, 2010). Digital CBT may help overcome such barriers because digital devices (e.g., computers, tablets, and smartphones) are widespread. Smartphone-based applications or “apps” are optimized and designed differently from computer-based digital CBT interventions and offer a promising delivery platform (Mohr et al., 2017). This is because they are owned by 81% of Americans (Pew Research Center, 2019) and permit real-time access to reliable, evidence-based interventions any time or place (Stolz et al., 2018). Smartphone-delivered interventions for anxiety were evaluated in a meta-analysis of 29 randomized controlled trials (RCTs). Results found significant reductions in GAD symptoms, with a small effect size compared with waitlist and/or inactive controls (Linnard, Cuijpers, Carlbring, Messer, & Fuller-Tyszkiewicz, 2019). Given that effect sizes for in-person CBT for GAD are typically large (Cuijpers et al., 2014), there is room for improvement in the efficacy of smartphone-delivered digital interventions.

The present intervention, *Daylight*[™] (<https://www.bighealth.com/daylight>), is a novel, smartphone-based and fully automated digital CBT intervention designed to enhance efficient learning and real-time application of CBT concepts and skills. For the goal of delivering engaging and digestible CBT techniques in a flexible, personalized manner, *Daylight* was developed in collaboration with designers, filmmakers, podcast producers, and animators. The structure, content, and design were all carefully considered to facilitate quick comprehension, frequent practice, and easy integration of CBT skills into the user's daily life.

Daylight was evaluated in a randomized, multiple-baseline single-case experimental design study of 21 participants (Miller et al., in review). Findings supported the feasibility, safety, and preliminary efficacy of *Daylight* at 6 weeks, though an RCT is necessary to test efficacy at the next level of rigor. We present a parallel-group superiority RCT examining digital CBT compared with a waitlist control in individuals with GAD. Our primary hypothesis was that digital CBT would be superior to waitlist control for reducing symptoms of GAD at post-intervention (6 weeks from randomization). Secondary aims evaluated effects for secondary outcomes, including worry, depressive symptoms, sleep difficulty, wellbeing, and participant-specific quality of life at post-

intervention and evaluated effects for both primary and secondary outcomes at follow-up (10 weeks). Further analyses adjusted for *Daylight* program use, examined whether effects on GAD symptoms were mediated by reductions in worry, depressive symptoms, or sleep difficulty and if baseline variables moderated effects.

2 | MATERIALS AND METHODS

2.1 | Design

A phase-II, randomized, partially blind parallel-group superiority trial with a primary endpoint at 6 weeks from randomization was conducted online and evaluated digital CBT for GAD (digital therapy for anxiety) compared with waitlist control. The trial design and participant flow are summarized in Figure 1. The trial was prospectively registered (ISRCTN12765810: <http://www.isrctn.com/ISRCTN12765810>) and ethical approval was obtained from the University of Oxford Medical Sciences Interdivisional Ethics Committee (MS-IDREC-C1-R61262/RE002). Apart from telephone calls, all aspects of the study, including screening, consent, assessment allocation, and delivery of the intervention were conducted online. The protocol has been published separately (Gu et al., 2020).

2.2 | Participants

Participants were 18 years or older with a diagnosis of GAD, as assessed by a score of ≥ 10 on the GAD-7 (Spitzer, Kroenke, Williams, & Löwe, 2006), combined with a positive screen for a GAD diagnosis on a digital version of the Mini-International Neuropsychiatric Interview (MINI) version-7 for DSM-5 (Sheehan et al., 2014). This self-report digital version included the same questions used in the standard MINI for GAD and did not include any additional follow-up questions. Individuals on prescription medication for anxiety, depressive symptoms, or poor sleep were included in the trial if they were on a stable dose for ≥ 4 weeks before baseline and were not allowed to have received CBT for anxiety in the last 12 months. Exclusion criteria included: self-reported diagnosis of schizophrenia, psychosis, bipolar disorder, seizure disorder, substance use disorder; recent trauma to the head or brain damage; severe cognitive impairment; serious physical health concerns necessitating surgery or with a prognosis of less than 6 months; or pregnancy. After initial online screening, participants underwent a telephone call with a research assistant (previously trained in the use of the MINI and supervised by a clinical psychologist who was available to help with any queries during the study) before consent to explain study procedures, answer any questions, and verify eligibility relating to GAD diagnosis with the Structured Clinical Interview for DSM-5 (SCID-5-CV; First, Williams, Karg, & Spitzer, 2015). Participants were recruited online through social media advertising and compensated in Amazon vouchers for completing each assessment at levels in accordance with the University of Oxford ethical requirements. They received £10 (\$12.30) for each assessment, plus a £10 bonus if all three were completed. Participants could receive a total of £40 (\$49.20) for completion of all

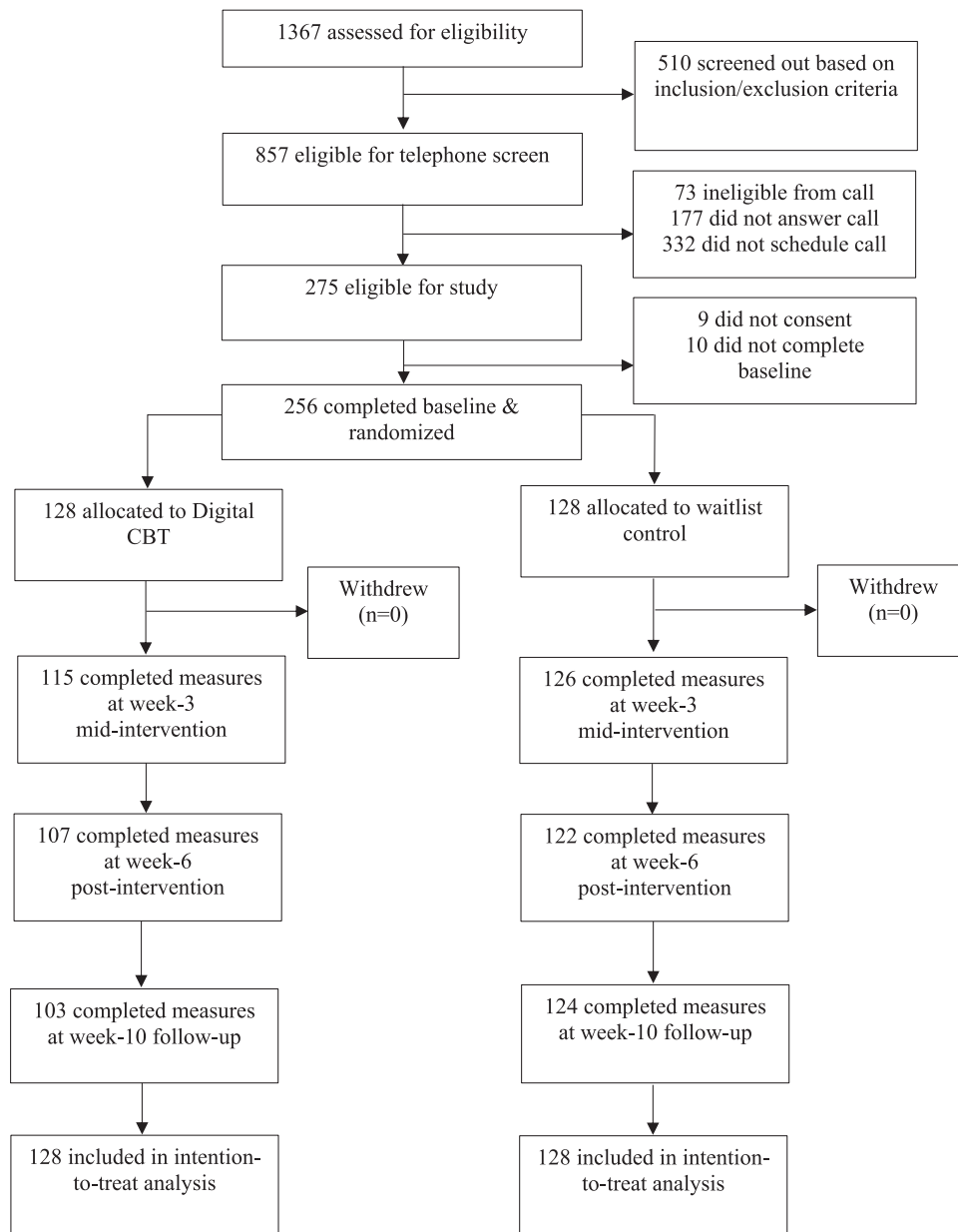


FIGURE 1 Participant flow through the study. CBT, cognitive behavioral therapy

assessments. Participants received verbal and written descriptions of study procedures and gave informed consent.

2.3 | Intervention

Participants in the digital CBT group were given access to *Daylight*, a fully automated personalized digital CBT program (iOS and Android release 2019). A virtual therapist guides individuals through the experience, which includes interactive exercises and animations to facilitate learning and implementation of CBT techniques. The four modules (~10–20 min in length) are based on principles of applied relaxation, stimulus control, cognitive restructuring, and imaginal

exposure (see Table ST1 for a description of each module). Techniques were selected and ordered based on (a) their effectiveness for GAD symptoms, (b) their effectiveness for symptoms often comorbid with GAD (e.g., sleep, low mood), and (c) their potential for translation into an engaging digital format. The program is designed to be self-paced. The app encourages daily use (e.g., practicing techniques in the app) and real-world implementation (e.g., practicing techniques in their daily lives, outside of the app). *Daylight* users receive reminders and encouragement to use the program in the form of emails, push notifications, and text messages. Users can opt-in to receive emails automatically upon sign-up and are asked whether they would like to opt-in to receive additional communications in the form of text messages and/or push notifications. Modules can

be repeated or shorter practice exercises (~5 min) can be accessed. Users are asked to complete weekly in-app brief assessments of anxiety, depressive symptoms, and sleep. Both the initial and weekly assessments and each exercise include elements of personalization. During assessments, users are provided with personalized feedback based on their self-reported anxiety, mood, and sleep, as well as their progress (e.g., users experiencing problems with sleep could receive a suggestion to practice a relaxation exercise before bedtime). The feedback provided in each exercise (e.g., providing additional instructions, troubleshooting, and/or guidance for future practice of the exercise) is tailored based on user inputs during the exercise (e.g., changes in their anxiety level during the exercise, whether they experienced any difficulties doing the exercise). Waitlist control participants did not receive any program during the study and were provided access to *Daylight* after the final follow-up assessment. Access to *Daylight* was not withdrawn and all participants were permitted to continue any other pharmacological or non-pharmacological interventions for GAD and use was tracked.

2.4 | Outcomes

Detailed protocol information is available elsewhere (Gu et al., 2020). Briefly, assessments were captured online using Qualtrics (Qualtrics, 2019) and measured at baseline (Week 0), mid-intervention (Week 3), post-intervention (Week 6), and follow-up (Week 10). Those randomized to the intervention arm only were invited to complete a longer-term uncontrolled assessment at Week 26 from randomization without compensation. The primary outcome was anxiety symptoms, measured by the GAD-7 (Spitzer et al., 2006) at post-intervention. Secondary outcomes included worry (Penn State Worry Questionnaire, PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990), depressive symptoms (Patient Health Questionnaire, PHQ-9; Kroenke, Spitzer, & Williams, 2001), sleep difficulty (Sleep Condition Indicator, SCI-8; Espie et al., 2014), and wellbeing (Warwick-Edinburgh Mental Wellbeing Scale, WEMWBS; Tennant et al., 2007). Participants completed an individualized quality of life measure (Patient-Generated Index; Ruta, Garratt, Leng, Russel, & MacDonald, 1994), which asks participants to rate their top three problem areas (e.g., relationships, work, etc.) associated with GAD. For anxiety (GAD-7), remission was assessed between baseline and post-intervention and baseline and follow-up if participants scored <10 (Spitzer et al., 2006) and reliable remission if participants scored <10 and demonstrated a change score of ≥ 5 , which is greater than the known unreliability of the measure (Richards & Borglin, 2011). Participants also demonstrated reliable deterioration if their GAD-7 scores increased by the reliable change score (≥ 5). Concomitant prescription or over-the-counter medication, and use of psychological therapies were documented at all assessment points. Participants' perceptions of intervention credibility (Deville & Borkovec, 2000) were recorded before randomization after being informed about the study conditions by telephone. Safety was assessed by the occurrence of any adverse events throughout the study period, reported spontaneously or in response to open-ended questions from study consent until the follow-up assessment. Using a modified Symptom Checklist (Kyle,

Morgan, Spiegelhalter, & Espie, 2011) at post-intervention, participants were asked to rate the occurrence of pre-specified unwanted symptoms that may have occurred (e.g., low mood, headache, fatigue, etc.) over the last 6 weeks.

2.5 | Sample size

A sample size of 242 participants (121 per group) was required to detect a between-group effect size of 0.5 with 90% power and a significance level of $p = .05$, accounting for 30% attrition (Gu et al., 2020). Due to a preponderance of females, recruitment was increased to include more people who identify as male and allow for a more representative sample of the GAD population (McLean, Asnaani, Litz, & Hofmann, 2011). In total, 256 participants were randomized.

2.6 | Randomization and blinding

Participants were randomized (1:1 allocation ratio) using a blocked randomization sequence online from Qualtrics survey software (Qualtrics, 2019). The study coordinator enrolled participants after the telephone call. Randomization and allocation were automated and carried out independently from the study team by Qualtrics. Participants and the trial coordinator were not blind to group allocation (digital CBT or waitlist control). The coordinator monitored uptake (download) of the intervention. All other members of the research team were blind to allocation. An external statistician (RE) was blind during the study and subgroup-unblind (groups labeled as "A" and "B") during the statistical analysis.

2.7 | Statistical methods

Analyses were intention-to-treat using Stata Version 16 (Stata Corp., 2017). The primary endpoint was analyzed using a linear mixed-effects model fitted to data at all post-randomization time points. The baseline outcome measure, group assignment, time, and time and group interactions were all included as fixed effects and participants were included as random intercepts to account for repeated measures. Missing outcome data were assumed to be missing at random.

Secondary hypotheses were tested using analogous analyses. Analyses of binary outcomes used logistic models on participants with outcome data to examine remission from initial assessment to post-intervention and initial assessment to follow-up, adjusting for baseline outcome measures. Sensitivity analyses tested the primary GAD-7 hypothesis using data from participants who downloaded the app, completed at least three or more modules, and completed all four modules of digital CBT. Separate sets of analyses were conducted for each of these three categories of intervention completion. For magnitude of change, we report Cohen's d standardized effect sizes (ES), estimated by dividing the adjusted mean difference by the pooled standard deviation (SD) of the corresponding outcome at baseline. A paired t test evaluated uncontrolled effects for outcomes for the

Daylight group only between baseline and Week 26 follow-up. Exploratory analyses examined potential mediators underlying the effect of digital CBT compared with waitlist control on GAD symptom severity in single mediator models and moderation analyses examined whether the between-group effect on GAD symptom severity was moderated by baseline variables (see Supporting Information for details). An unplanned interim analysis for efficacy was performed for post-intervention after collection of all participant data and partially unblinded the statistician for the follow-up assessment only.

3 | RESULTS

The sample comprised 256 adults randomized to either digital CBT ($n = 128$) or waitlist control ($n = 128$), of whom 175 (68%) were female. The mean (SD) age was 30.9 (10.7), with a range of 18–67 years. Most participants were from the United States ($n = 193$), and 63 were from the UK. Table 1 provides an overview of demographics and baseline scores for both primary and secondary outcomes by group. All 256 participants were recruited between August 2 and November 7, 2019, and the final controlled (Week 10) and uncontrolled (Week 26) follow-up assessments were completed on January 15 and May 11, 2020, respectively. No participants withdrew from the study.

In the digital CBT group, 127/128 (99%) downloaded and initiated the intervention, 121 (95%) completed at least 1 module, 108 (84%) completed at least 2 modules, 92 (72%) completed at least 3 modules, and 76 (59%) completed all 4 modules. On average, participants completed 3.10 ($SD = 1.27$) of 4 modules and 14.55 ($SD = 20.47$; range of 0–193) pieces of content. Content counts include the number of times modules (which could be repeated) were completed, as well as the number of practice versions of modules completed. Ninety-one (71%) of digital CBT participants opted to receive engagement text messages, and 119 (93%) opted to receive push notifications.

3.1 | Intervention effects on the primary outcome

Table 2 presents the results (unadjusted, adjusted differences, and Cohen's d), comparing differences between digital CBT and waitlist control over time (mid- and post-intervention, and follow-up). There was a significant reduction in GAD-7 anxiety scores in the digital CBT group compared with the waitlist control group at all assessment time points (Figure 2), with large between-group effect size scores at mid-intervention ($d = 0.96$), post-intervention (primary endpoint: $d = 1.08$), and at follow-up ($d = 1.43$). Sensitivity analyses with different levels of digital CBT program use found similar results as reported here (see Tables ST2 and ST3).

3.2 | Intervention effects on secondary outcomes

Digital CBT led to significant improvements compared with waitlist control for worry (PSWQ), depressive symptoms (PHQ-9),

TABLE 1 Descriptive characteristics at baseline

Characteristic	Waitlist control ($n = 128$)	Digital CBT ($n = 128$)
Age, mean (SD), range (years)	30.7 (10.6), 18–67	31.1 (10.8), 18–67
Country, n (%)		
UK	33 (26%)	30 (23%)
United States	95 (74%)	98 (77%)
Sex, n (%)		
Female	89 (70%)	86 (67%)
Male	36 (28%)	41 (32%)
Other	3 (2%)	1 (1%)
Ethnic origin, n (%)		
Arab	1 (1%)	0 (0%)
Asian	5 (4%)	2 (2%)
Black	5 (4%)	5 (4%)
Mixed	5 (4%)	8 (6%)
Other	6 (5%)	3 (2%)
White	106 (83%)	109 (85%)
Did not state	0 (0%)	1 (1%)
Education, n (%)		
No formal qualifications	2 (2%)	3 (2%)
Secondary school/high school graduate	20 (16%)	19 (15%)
College	31 (24%)	35 (27%)
Undergraduate/bachelor's degree	50 (39%)	42 (33%)
Postgraduate or professional degree	25 (20%)	29 (23%)
Relationship status, n (%)		
Married/civil partnership	35 (27%)	33 (26%)
Separated	3 (2%)	1 (1%)
Divorced	7 (6%)	8 (6%)
Never married/single	62 (48%)	61 (48%)
Partnered	20 (16%)	24 (19%)
Prefer not to say	1 (1%)	1 (1%)
Duration of anxiety and worry		
6 months to 1 year	7 (6%)	1 (0.8%)
1–3 years	15 (12%)	19 (15%)
Over 3 years	106 (83%)	107 (84%)
Days taken anxiety medication prescribed or not in the last 3 weeks, mean (SD)	6.7 (9.5)	5.2 (9.0)
Number of individuals who reported using anxiety medication in the last 3 weeks (%)	49 (38%)	36 (28%)
Days seen a treatment provider for anxiety in the last 3 weeks, mean (SD)	0.6 (2.1)	0.5 (1.1)
Number of individuals who saw a treatment provider for anxiety in the last 3 weeks (%)	34 (26%)	31 (24%)

(Continues)

TABLE 1 (Continued)

Characteristic	Waitlist control (n = 128)	Digital CBT (n = 128)
Outcomes at baseline, mean (SD) ^a		
Anxiety symptoms, (GAD-7)	15.6 (3.1)	15.6 (2.8)
Worry symptoms (PSWQ)	70.9 (5.9)	70.8 (6.9)
Depressive symptoms (PHQ-9)	14.7 (5.8)	15.4 (5.1)
Sleep difficulty (SCI-8)	11.7 (6.6)	12.0 (5.5)
Mental wellbeing (WEMWBS)	34.4 (8.0)	34.4 (7.7)
Patient-Generated Index		
Most important concern	38.7 (21.6)	37.5 (22.7)
Second most important concern	40.3 (21.9)	39.2 (19.1)
Third most important concern	41.9 (21.2)	40.7 (22.0)
Total credibility (CEQ)	19.7 (3.9)	19.2 (4.1)

Abbreviations: CBT, cognitive behavioral therapy; CEQ, Credibility and Expectancy Questionnaire; GAD-7, generalized anxiety disorder; PHQ-9, 9-item Patient Health Questionnaire; PSWQ, Penn State Worry Questionnaire; SCI-8, 8-item Sleep Condition Indicator; SD, standard deviation; WEMWBS, Warwick-Edinburgh Mental Wellbeing Scale.

^aOutcome assessment scales are explained in Section 2.

sleep difficulty (SCI-8), and wellbeing (WEMWBS) at all assessment time points (Table 3). The effect size scores ranged from moderate-to-large ($d > .51$) for all except the SCI-8, which was small at both mid- ($d = .24$) and post-intervention ($d = .31$), and moderate at follow-up ($d = 0.65$) (Figure 3). Improvements were maintained at the uncontrolled Week 26 follow-up for those who received the intervention (see Table ST4). Significant improvements were found for digital CBT participants' first most important area for concern on the Patient-Generated Index with small-to-moderate effect sizes at post-intervention ($d = .34$) and follow-up ($d = .41$). No significant differences were present for the second most important concern at post-intervention; however, significant differences were observed at follow-up with a moderate effect ($d = .42$). Significant differences were present for participants' third most important concern at both post-intervention ($d = .29$) and follow-up ($d = .67$) with small and moderate effect sizes, respectively.

3.3 | Remission and reliable remission

At post-intervention (Week 6), participants in the digital CBT group were more likely to experience remission of anxiety (65/107 [61%]) compared with waitlist control (38/122 [31%]) (OR [SE] = 3.99 [1.20]; 95% CI = 2.21, 7.21; $p < .001$). Similarly, 56/107 (52%) in the digital CBT group versus 33/122 (27%) in the waitlist control showed reliable remission in anxiety at post-intervention (OR [SE] = 3.05 [0.87]; 95% CI = 1.74, 5.34; $p < .001$). At follow-up (Week 10), 73/103 (71%) in the digital CBT group experienced remission compared with 41/124 (33%) in the waitlist control (OR [SE] = 5.84 [1.84]; 95% CI = 3.15, 10.82; $p < .001$) and 66/103 (64%) in the digital CBT group showed reliable remission compared with 34/124 (27%) in the waitlist control (OR [SE] = 4.84 [1.41]; 95% CI = 2.73, 8.57; $p < .001$). The denominators are lower than the full sample due to missing data at post- and follow-up. Due to the assumption that data are missing at random, we believe these percentages to be largely representative of the full intention-to-treat sample. Table ST5 reports effects for anxiety, depressive symptoms, and sleep difficulty. In terms of reliable deterioration (increase in GAD-7 scores of ≥ 5) in the digital CBT group, three participants experienced this at post-intervention and two at follow-up. In the control group, four experienced reliable deterioration at post-intervention and seven at follow-up.

3.4 | Safety

One participant only, in the digital CBT group, reported experiencing an adverse event related to difficulty accessing the intervention (signing up for *Daylight* and downloading the app), which the participant deemed as distressing and as contributing to increased anxiety. This event was reported to the Ethics Committee and the participant opted not to use the intervention but remained in the trial by continuing with the study assessments only. Participants in the digital CBT group reported significantly fewer occurrences of unwanted symptoms from the Symptom Checklist than the control group during the 6-week study period, including low mood, fatigue and/or exhaustion, extreme sleepiness, feeling agitated, difficulty remembering things, headache and/or migraine, difficulty

TABLE 2 Effects of digital cognitive behavioral therapy versus waitlist control on the primary outcome

Assessment ^a	Unadjusted, mean (SD)		Adjusted difference (SE) (95% CI)	Cohen's <i>d</i>	<i>p</i> value
	Waitlist control (<i>n</i> = 128)	Digital CBT (<i>n</i> = 128)			
GAD-7					
Baseline (Week 0)	15.6 (3.1)	15.6 (2.8)			
Mid-intervention (Week 3)	13.64 (4.43)	10.81 (4.39)	−2.85 (0.54) (−3.91, −1.78)	0.96 (0.60, 1.31)	<0.001
Post-intervention (Week 6)	12.66 (4.70)	9.46 (4.67)	−3.22 (0.55) (−4.31, −2.14)	1.08 (0.72, 1.45)	<0.001
Follow-up (Week 10)	12.15 (4.73)	7.90 (4.85)	−4.26 (0.56) (−5.35, −3.17)	1.43 (1.07, 1.80)	<0.001

Abbreviations: CI, confidence interval; GAD-7, 7-item generalized anxiety disorder; SD, standard deviation; SE, standard error.

^aOutcome assessment scales are explained in Section 2.

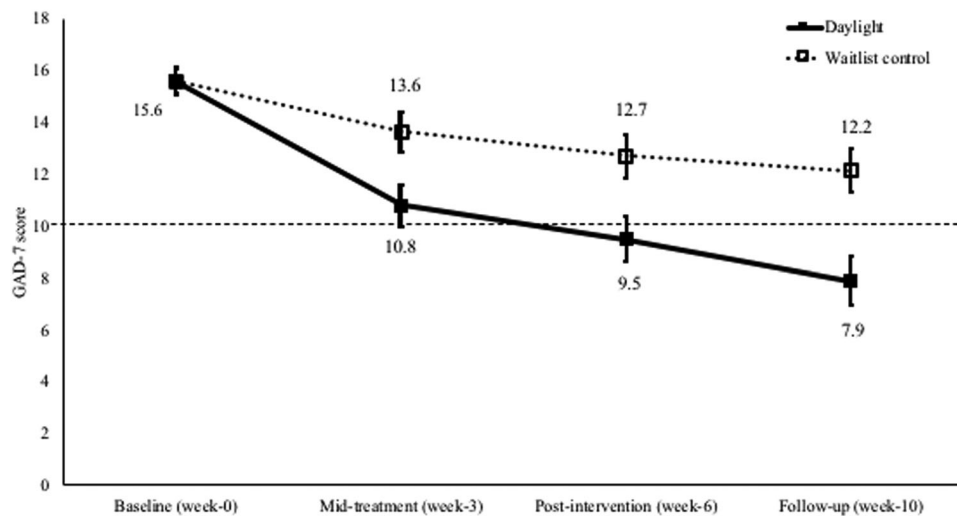


FIGURE 2 Global anxiety symptoms between groups over time. Average symptoms of anxiety measured by the 7-item generalized anxiety disorder questionnaire (GAD-7) assessed at baseline (Week 0; $n = 256$), mid-intervention (Week 3; $n = 241$), post-intervention (Week 6; $n = 229$), and follow-up (Week 10; $n = 227$). Error bars represent 95% confidence intervals. The smaller, dotted line illustrates the clinical cut-off for anxiety (score of 10 or higher)

concentrating and focusing on things, reduced motivation and/or energy, blurred vision, dizziness, and feeling irritable (see Table ST6).

3.5 | Exploratory mediation and moderation analyses

Post-intervention effects on the GAD-7 were mediated by mid-intervention measures of worry (PSWQ): proportion mediated = 21%, depressive symptoms (PHQ-9): 25%, and sleep difficulty (SCI-8): 13% (see Table ST7). No baseline variables moderated the effects of the intervention on the GAD-7 (see Table ST8).

4 | DISCUSSION

This trial is the first between-group efficacy and safety evaluation of a novel smartphone-delivered digital CBT intervention, *Daylight*, in participants with GAD. *Daylight* was designed to enhance efficient learning and implementation of CBT concepts and skills with engaging and brief techniques, personalized content and guidance, and flexible pacing. Results supported the primary hypothesis, that this fully automated intervention leads to large improvements in anxiety symptoms compared with a waitlist control at post-intervention (6 weeks; $d = 1.08$) and follow-up (10 weeks; $d = 1.43$). Results are noteworthy given that average baseline GAD-7 scores for both groups were in the severe range (mean > 15; Spitzer et al., 2006). Improvements were clinically meaningful as 52% of participants in the digital CBT group experienced reliable remission in anxiety symptoms at post-intervention compared with 27% in the waitlist control. This increased to 71% versus 27%, respectively, in participants with outcome data at follow-up. The odds ratio for a reliable

remission with the digital CBT intervention (compared with waitlist control) was 3.05 at post-intervention and increased to 5.84 at follow-up.

We found improvements in secondary outcomes, underscoring the potential benefits of digital CBT on broader aspects of mental health and functioning (Figure 3) compared with a waitlist control. Significant improvements were observed in worry, depressive symptoms, sleep difficulty, wellbeing, and participant-specific quality of life at both post-intervention ($d = 0.34$ – 0.73) and follow-up ($d = 0.43$ – 1.11). Despite moderate-to-large between-group differences for worry at post-intervention, the raw change within the digital CBT group was relatively small. The PSWQ was designed to evaluate trait worry and may be less sensitive to change (Meyer et al., 1990; Verkuil, Brosschot, & Thayer, 2007). Depressive symptom comorbidity was common in participants as expected in GAD (Kessler et al., 2005), and there were moderate-to-large effects on depressive symptoms at both post-intervention and follow-up. Longer-term (uncontrolled) improvements were maintained for the *Daylight* group across outcomes at Week 26 follow-up. This is notable given these data were collected between February and May 2020, after the onset of the COVID-19 pandemic.

The low number ($n = 1$) and mild nature of this adverse event, in combination with good engagement, suggest *Daylight* is safe and well tolerated. Additionally, the frequency of unwanted symptoms (e.g., agitation, low mood, fatigue/exhaustion) occurring during the study, as reported on the Symptom Checklist, was significantly higher in the waitlist control. Greater unwanted symptoms reported in the waitlist group may reflect general untreated symptoms of GAD and related comorbid symptoms (such as depressive symptoms) that improved only in the digital CBT group. The intervention was credible with moderate-high scores on average, acceptable (99% downloaded the app), and engaging (72% and 59% completed 3 of 4 and all 4 modules,

TABLE 3 Effects of digital cognitive behavioral therapy versus waitlist control on secondary outcomes

	Unadjusted, mean (SD)				
Assessment ^a	Waitlist control (n = 128)	Digital CBT (n = 128)	Adjusted difference (SE) (95% CI)	Cohen's <i>d</i>	<i>P</i> value
PSWQ					
Baseline (Week 0)	70.9 (5.9)	70.8 (6.9)			
Mid-intervention (Week 3)	68.38 (7.29)	64.35 (9.26)	−3.97 (0.98) (−5.90, −2.04)	0.62 (0.32, 0.92)	<.001
Post-intervention (Week 6)	67.66 (7.84)	62.56 (9.81)	−4.68 (1.00) (−6.65, −2.72)	0.73 (0.42, 1.04)	<.001
Follow-up (Week 10)	67.38 (7.98)	59.95 (11.19)	−7.09 (1.00) (−9.06, −5.11)	1.11 (0.80, 1.42)	<.001
PHQ-9					
Baseline (Week 0)	14.7 (5.8)	15.4 (5.1)			
Mid-intervention (Week 3)	13.99 (5.88)	12.19 (6.30)	−2.22 (0.64) (−3.48, −0.97)	0.41 (0.18, 0.64)	<.001
Post-intervention (Week 6)	13.15 (6.15)	10.80 (6.00)	−2.80 (0.65) (−4.07, −1.52)	0.51 (0.28, 0.75)	<.001
Follow-up (Week 10)	13.66 (6.79)	9.37 (6.21)	−4.43 (0.65) (−5.71, −3.15)	0.81 (0.58, 1.05)	<.001
SCI-8					
Baseline (Week 0)	11.7 (6.6)	12.0 (5.5)			
Mid-intervention (Week 3)	13.99 (8.06)	14.97 (7.35)	1.69 (0.68) (0.35, 3.03)	0.24 (0.05, 0.43)	.013
Post-intervention (Week 4)	12.93 (7.06)	14.68 (6.38)	2.18 (0.70) (0.82, 3.55)	0.31 (0.12, 0.50)	.002
Follow-up (Week 10)	13.90 (8.35)	17.93 (7.58)	4.57 (0.70) (3.20, 5.94)	0.65 (0.45, 0.84)	<.001
WEMWBS					
Baseline (Week 0)	34.4 (8.0)	34.4 (8.0)			
Post-intervention (Week 6)	35.68 (9.36)	40.68 (9.50)	5.04 (1.04) (3.00, 7.08)	0.64 (0.38, 0.90)	<.001
Follow-up (Week 10)	35.31 (8.80)	43.43 (10.18)	7.94 (1.05) (5.89, 9.99)	1.01 (0.75, 1.27)	<.001
Patient-Generated Index					
Most important concern					
Baseline (Week 0)	38.7 (21.6)	37.5 (22.7)			
Post-intervention (Week 6)	46.33 (22.15)	52.17 (25.75)	7.44 (2.90) (1.75, 13.13)	0.34 (0.08, 0.59)	.010
Follow-up (Week 10)	48.10 (23.50)	58.02 (24.85)	9.61 (2.93) (3.88, 15.34)	0.43 (0.18, 0.69)	.001
Second most important concern					
Baseline (Week 0)	40.3 (21.9)	39.2 (19.1)			
Post-intervention (Week 6)	49.66 (23.53)	52.72 (23.44)	4.10 (3.02) (−1.80, 10.01)	0.20 (−0.09, 0.49)	.173
Follow-up (Week 10)	46.28 (22.88)	53.70 (24.52)	8.58 (3.03) (2.65, 14.51)	0.42 (0.13, 0.71)	.005
Third most important concern					
Baseline (Week 0)	41.9 (21.2)	40.7 (22.0)			
Post-intervention (Week 6)	46.39 (21.46)	51.92 (23.40)	6.25 (2.97) (0.42, 12.08)	0.29 (0.02, 0.56)	.036
Follow-up (Week 10)	43.83 (21.70)	58.20 (24.92)	14.49 (2.99) (8.63, 20.36)	0.67 (0.40, 0.94)	<.001

Abbreviations: CI, confidence interval; PHQ-9, 9-item Patient Health Questionnaire; PSWQ, Penn State Worry Questionnaire; SCI-8, 8-item Sleep Condition Indicator; SD, standard deviation; SE, standard error; WEMWBS, Warwick-Edinburgh Mental Wellbeing Scale.

^aOutcome assessment scales are explained in Section 2.

respectively). For the primary outcome (GAD-7), sensitivity analyses in those who used Daylight found larger adjusted differences with greater module use (see Tables ST2 and ST3).

Exploratory mediation analyses found reductions in anxiety associated with the digital CBT intervention were partially explained by improvements in worry, depressive symptoms, and sleep when entered individually. Worry is a direct target of the intervention, thus a logical mediator. Depressive symptoms and sleep difficulty may have functioned as mediators due to participants applying program exercises to depressive or sleep-related cognitions or due to natural bidirectional interactions in comorbid symptoms (Jansson-Fröjmark & Lindblom, 2008). Baseline variables, including demographics (e.g., sex, age),

concomitant treatment use, and symptoms of anxiety, worry, and depression, did not moderate the effect of digital CBT on GAD-7 scores at post-intervention, supporting the generalizability of these data.

Findings are consistent with our recent open trial which examined the feasibility, safety, and preliminary efficacy of *Daylight* (Miller et al., in review). Together, these studies provide evidence that *Daylight*, when used in conjunction with usual care, is an efficacious and safe nonpharmacological intervention for GAD. This aligns with previous research suggesting that smartphone-based digital CBT can be safe and effective; however, with *Daylight*, we see large effects on GAD symptoms, whereas published data on other smartphone-based digital CBT apps have suggested small effects (Linardon et al., 2019). It is

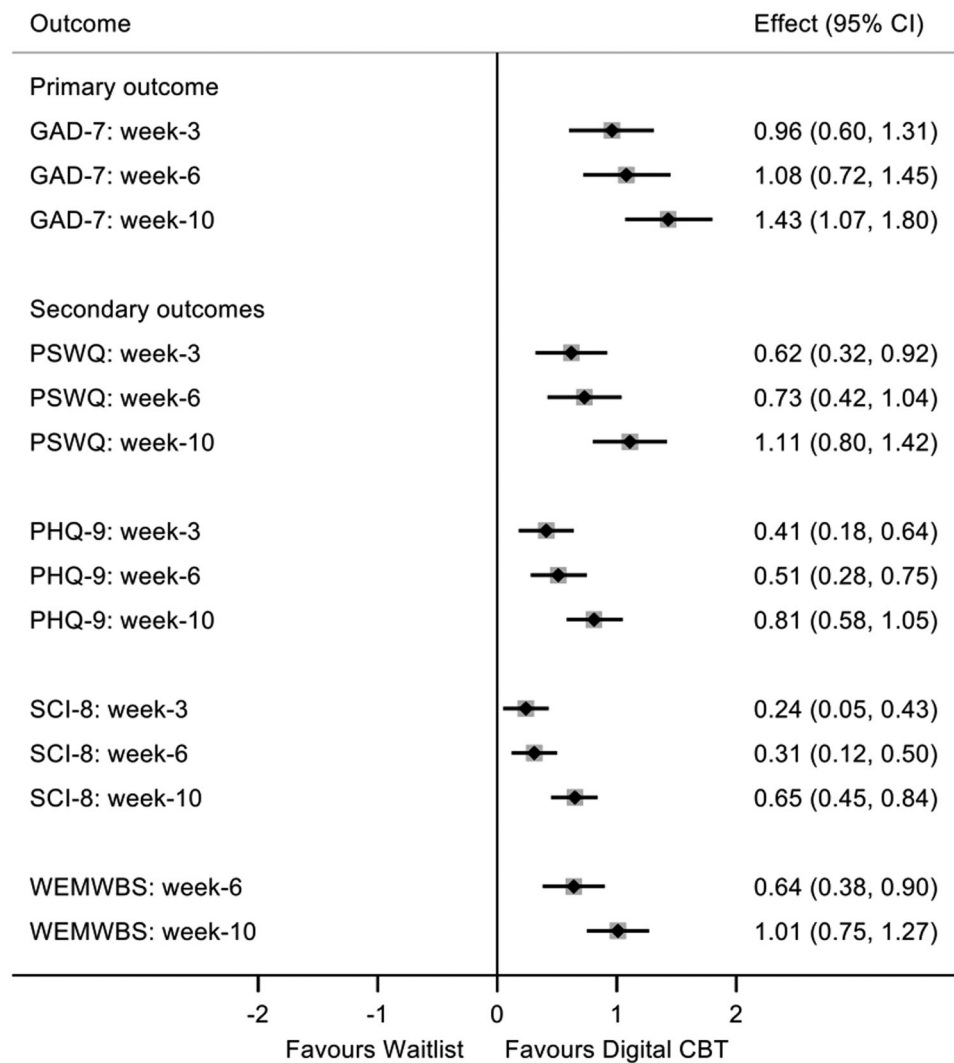


FIGURE 3 Between-group effect sizes (d) for primary and secondary outcomes at mid-treatment (Week 3), post-intervention (Week 6) and follow-up (Week 10). CBT, cognitive behavioral therapy; CI, confidence interval; GAD-7, generalized anxiety disorder; PHQ-9, 9-item patient health questionnaire; PSWQ, Penn State Worry Questionnaire; SCI-8, 8-item Sleep Condition Indicator; WEMWBS, Warwick-Edinburgh Mental Wellbeing Scale

possible that the novel design of *Daylight* (i.e., intended for flexible daily use; optimized for engagement and accelerated learning of effective techniques) led to these larger effects.

The study has limitations. First, due to our interest in the scalable potential of smartphone-based digital CBT for GAD, both our outcome measures and inclusion criteria relied on self-reported symptoms. We did, however, verify baseline GAD diagnosis with a digital version of the MINI for DSM-5 (Sheehan et al., 2014) and the Structured Clinical Interview for DSM-5 (First et al., 2015) in a telephone call. Second, findings may not be generalizable to more diverse populations, and future studies should evaluate this, explore moderator effects in larger samples, and consider durability of effects over the longer term. Further, for this study, we selected a waitlist as the suitable control condition for digital CBT; this control group provides an index of the efficacy of digital CBT as would be offered by a provider, insurance company, or employer. This does not,

however, control for potential expectation, demand, and attention effects. This study did not examine whether specific elements of digital CBT offered differential efficacy. Remission of GAD was pre-defined using a cut-off point of <10 on the GAD-7 and participants may have been classified as remitted but still experience subthreshold symptoms. Strengths of this study include the high rate of outcome assessment completion ($\geq 89\%$ across all controlled time points), good participant uptake, and subsequent digital CBT program use. The sample was also clinically relevant because participants reported previous use of pharmacologic or psychosocial treatment.

5 | SUMMARY

The results of this trial indicate that a novel smartphone-delivered, fully automated digital CBT intervention, *Daylight*, is efficacious and

safe for improving anxiety symptoms in adults with GAD compared with waitlist control. Anxiety symptoms improved over time, and other pertinent areas of mental health and functioning including worry, depressive symptoms, sleep difficulty, wellbeing, and participant-specific quality of life also showed significant gains. Our findings suggest that digital CBT can be efficacious for individuals with GAD and may prove to be a scalable solution that avoids some of the burdens and limitations associated with in-person CBT.

ACKNOWLEDGMENTS

We would like to thank all of the participants who were involved in this study. Slaton Z. Freeman and Annabelle F. DiVita were assistant project coordinators and assisted with recruitment and data collection. We would also like to thank the clinical, creative, product, and design teams involved in developing the digital CBT intervention Daylight. The research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). KEAS is supported by the NIHR Oxford Health BRC (Grant No. BRC-1215-20005). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

CONFLICT OF INTERESTS

C. A. E. is cofounder and Chief Medical Officer of Big Health Inc. is salaried and is a shareholder in the company. J. R. C. and A. L. H. are employed by Big Health Inc., are salaried and are shareholders in the company. O. S., C. B. M., R. S., and M. L. D. are employed by Big Health Inc. and are salaried by the company. R. E. is a statistical consultant for Big Health Inc. J. A. J. S. is a paid Clinical Advisor for Big Health Inc. and receives royalty payments from Oxford University Press and Academic Press. M. W. O. is a Scientific and Clinical Advisor for Big Health Inc. and shareholder, and has received speaker support. Remaining authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Concept and design: J. R. C., C. B. M., A. L. H., and C. A. E.; acquisition, analysis, and interpretation of data: all authors; critical revision of the manuscript for important intellectual content: all authors; statistical analysis: R. E.; administrative, technical, or participant support: A. L. H., M. L. D., O. S., and C. B. M.; study supervision: C. B. M., J. A. J. S., J. R. C., and C. A. E.

DATA AVAILABILITY STATEMENT

The full study protocol, statistical code and participant-level data set for the current study may be available from the corresponding author on reasonable request.

ORCID

Christopher B. Miller  <http://orcid.org/0000-0002-2936-7717>

Michelle G. Craske  <http://orcid.org/0000-0002-3704-5240>

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)* (5th ed.). Washington, DC: Author.

- Anxiety and Depression Association of America (2020). *Treatment*. Association of America. Retrieved from <https://adaa.org/living-with-anxiety/children/treatment>
- Canadian Psychiatric Association. (2006). Clinical practice guidelines. Management of anxiety disorders. *Canadian Journal of Psychiatry*. *Revue Canadienne de Psychiatrie*, 51(8 Suppl. 2), 9S–91S.
- Carl, E., Witcraft, S. M., Kauffman, B. Y., Gillespie, E. M., Becker, E. S., Cuijpers, P., ... Powers, M. B. (2020). Psychological and pharmacological treatments for Generalized Anxiety Disorder (GAD): A meta-analysis of randomized controlled trials. *Cognitive Behaviour Therapy*, 49(1), 1–21. <https://doi.org/10.1080/16506073.2018.1560358>
- Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A. J., & Hofmann, S. G. (2018). Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials. *Depression and Anxiety*, 35(6), 502–514. <https://doi.org/10.1002/da.22728>
- Comer, J. S., & Barlow, D. H. (2014). The occasional case against broad dissemination and implementation: Retaining a role for specialty care in the delivery of psychological treatments. *The American Psychologist*, 69(1), 1–18. <https://doi.org/10.1037/a0033582>
- Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., & Andersson, G. (2014). Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clinical Psychology Review*, 34(2), 130–140. <https://doi.org/10.1016/j.cpr.2014.01.002>
- Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31(2), 73–86. [https://doi.org/10.1016/s0005-7916\(00\)00012-4](https://doi.org/10.1016/s0005-7916(00)00012-4)
- Espie, C. A., Kyle, S. D., Hames, P., Gardani, M., Fleming, L., & Cape, J. (2014). The sleep condition indicator: A clinical screening tool to evaluate insomnia disorder. *BMJ Open*, 4(3), 004183. <https://doi.org/10.1136/bmjopen-2013-004183>
- First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2015). *Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV)* (pp. 1–94). Arlington, VA: American Psychiatric Association.
- Gu, J., Miller, C. B., Henry, A. L., Espie, C. A., Davis, M. L., Stott, R., ... Carl, J. R. (2020). Efficacy of digital cognitive behavioural therapy to treat symptoms of generalised anxiety disorder: A study protocol for a randomised controlled trial. *Trials*, 21, 357. <https://doi.org/10.1186/s13063-020-4230-6>
- Gunter, R. W., & Whittal, M. L. (2010). Dissemination of cognitive-behavioral treatments for anxiety disorders: Overcoming barriers and improving patient access. *Clinical Psychology Review*, 30(2), 194–202. <https://doi.org/10.1016/j.cpr.2009.11.001>
- Jansson-Fröjmark, M., & Lindblom, K. (2008). A bidirectional relationship between anxiety and depression, and insomnia? A prospective study in the general population. *Journal of Psychosomatic Research*, 64(4), 443–449. <https://doi.org/10.1016/j.jpsychores.2007.10.016>
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, 62(6), 593–602. <https://doi.org/10.1001/archpsyc.62.6.593>
- Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, 21(3), 169–184. <https://doi.org/10.1002/mpr.1359>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kyle, S. D., Morgan, K., Spiegelhalter, K., & Espie, C. A. (2011). No pain, no gain: An exploratory within-subjects mixed-methods evaluation of the

- patient experience of sleep restriction therapy (SRT) for insomnia. *Sleep Medicine*, 12(8), 735–747. <https://doi.org/10.1016/j.sleep.2011.03.016>
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry*, 18, 325–336. <https://doi.org/10.1002/wps.20673>
- Locke, A. B., Kirst, N., & Shultz, C. G. (2015). Diagnosis and management of generalized anxiety disorder and panic disorder in adults. *American Family Physician*, 91(9), 617–624.
- McLean, C. P., Asnaani, A., Litz, B. T., & Hofmann, S. G. (2011). Gender differences in anxiety disorders: Prevalence, course of illness, comorbidity and burden of illness. *Journal of Psychiatric Research*, 45(8), 1027–1035. <https://doi.org/10.1016/j.jpsychires.2011.03.006>
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28(6), 487–495. [https://doi.org/10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6)
- Miller, C. B., Gu, J., Henry, A. L., Davis, M. L., Espie, C. A., Stott, R., ... Carl, J. R. Feasibility and efficacy of digital cognitive behavioral therapy for generalized anxiety disorder symptoms: A randomized multiple-baseline study. *Journal of Consulting and Clinical Psychology*. (in review).
- Mohr, D. C., Tomasino, K. N., Lattie, E. G., Palac, H. L., Kwasny, M. J., Weingardt, K., & Caccamo, L. (2017). IntelliCare: An eclectic, skills-based app suite for the treatment of depression and anxiety. *Journal of Medical Internet Research*, 19(1), e10. <https://doi.org/10.2196/jmir.6645>
- National Collaborating Centre for Mental Health. (2018). The Improving Access to Psychological Therapies Manual Appendices and helpful resources. Retrieved from <https://www.england.nhs.uk/wp-content/uploads/2018/06/iapt-manual-resources-v2.pdf>
- National Institute for Health and Clinical Excellence. (2011). Clinical Guideline [CG113] Generalised anxiety disorder and panic disorder in adults: Management.
- Pew Research Center. (2019). *Demographics of mobile device ownership and adoption in the United States*. Retrieved from <https://www.pewresearch.org/internet/fact-sheet/mobile/>
- Qualtrics. (2019). Qualtrics, Version August 2019. Provo, UT. Retrieved from <http://www.qualtrics.com>
- Revicki, D. A., Travers, K., Wyrwich, K. W., Svedsäter, H., Locklear, J., Matterna, M. S., ... Montgomery, S. (2012). Humanistic and economic burden of generalized anxiety disorder in North America and Europe. *Journal of Affective Disorders*, 140(2), 103–112. <https://doi.org/10.1016/j.jad.2011.11.014>
- Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: Two year prospective cohort study. *Journal of Affective Disorders*, 133(1–2), 51–60. <https://doi.org/10.1016/j.jad.2011.03.024>
- Ruta, D. A., Garratt, A. M., Leng, M., Russell, I. T., & MacDonald, L. M. (1994). A new approach to the measurement of quality of life. The Patient-Generated Index. *Medical Care*, 32(11), 1109–1126. <https://doi.org/10.1097/00005650-199411000-00004>
- Sheehan, D., Janavs, J., Baker, R., Sheehan, K.H., Knapp, E., & Sheehan, M. (2014). *Mini International Neuropsychiatric Interview—Version 7.0.0 DSM-5*.
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: Author.
- Stolz, T., Schulz, A., Krieger, T., Vincent, A., Urech, A., Moser, C., & Berger, T. (2018). A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. *Journal of Consulting and Clinical Psychology*, 86(6), 493–504. <https://doi.org/10.1037/ccp0000301>
- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., ... Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*, 5(1), 63. <https://doi.org/10.1186/1477-7525-5-63>
- Verkuil, B., Brosschot, J. F., & Thayer, J. F. (2007). Capturing worry in daily life: Are trait questionnaires sufficient? *Behaviour Research and Therapy*, 45(8), 1835–1844. <https://doi.org/10.1016/j.brat.2007.02.004>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Carl JR, Miller CB, Henry AL, et al. Efficacy of digital cognitive behavioral therapy for moderate-to-severe symptoms of generalized anxiety disorder: A randomized controlled trial. *Depression and Anxiety*. 2020;1–11. <https://doi.org/10.1002/da.23079>