Journal of Behavior Therapy and Experimental Psychiatry 70 (2021) 101609

Contents lists available at ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep

Feasibility and efficacy of a digital CBT intervention for symptoms of Generalized Anxiety Disorder: A randomized multiple-baseline study☆

Christopher B. Miller[a,b,c,*], Jenny Gu[a,b,d], Alasdair L. Henry[a,b,c], Michelle L. Davis[a,b], Colin A. Espie[a,b,c], Richard Stott[a,b,d], Adrienne J. Heinz[a,b], Kate H. Bentley[f], Guy M. Goodwin[d], Bernard S. Gorman[e], Michelle G. Craske[g], Jenna R. Carl[a,b]

[a] Big Health Inc., San Francisco, CA, USA
[b] Big Health Inc., London, UK
[c] Sleep and Circadian Neuroscience Institute, Nuffield Department of Clinical Neurosciences, Sir William Dunn School of Pathology, University of Oxford, Oxford, UK
[d] University Department of Psychiatry, University of Oxford and Oxford Health NHS Trust, Warneford Hospital, Oxford, UK
[e] Department of Psychology, Harvard University, Cambridge, MA, USA
[f] Derner School of Psychology, Adelphi University, Garden City, NY, USA
[g] Anxiety and Depression Research Centre (ADRC), UCLA, Los Angeles, CA, USA

ARTICLE INFO

Keywords:
Anxiety
Digital
Cognitive behavioral therapy
Single-case experimental design
Smartphone

ABSTRACT

Background and objectives: Cognitive behavioral therapy (CBT) is a first-line treatment for anxiety, but it is not widely available as clinical guidelines recommend. We examined the feasibility and efficacy of a novel smartphone-based fully automated digital CBT intervention, '*Daylight*™', to improve symptoms of Generalized Anxiety Disorder (GAD).

Methods: In this multiple-baseline design, 21 adults (20 F; mean age 43yrs. range 19–65yrs.) with moderate-to-severe symptoms of GAD were randomized to one of three baseline durations (2-, 4-, or 6-weeks) and then received access to digital CBT. Participants completed daily ratings of anxiety and worry, weekly measures of anxiety, depressive symptoms, and sleep, and measures of anxiety, worry, wellbeing, quality of life, CBT skill acquisition, and work performance at initial assessment prior to baseline randomization, post-intervention, and follow-up.

Results: Digital CBT was found to be feasible in terms of engagement, satisfaction, and safety. For preliminary efficacy, improvements were detected in daily and weekly outcomes of anxiety for most participants. Despite individual differences, significant improvements occurred with the introduction of digital CBT and not during baseline. Overall, 70% of participants no longer had clinically significant symptoms of GAD, 61% no longer had significant depressive symptoms, and 40% no longer had significant sleep difficulty at post-intervention.

Limitations: The study sample was recruited using the internet and was mostly female, limiting the generalizability of the findings.

Conclusions: Findings support the feasibility and efficacy of *Daylight*. Further examination in randomized controlled trials is now warranted.

1. Introduction

Generalized Anxiety Disorder (GAD) is characterized by symptoms of excessive worry and anxiety that are difficult to control (American Psychiatric Association, 2013). Approximately 5–8% of adults are affected by GAD (Kessler et al., 2005; Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007; Roy-Byrne & Wagner, 2004), effects of which include impaired health status, wellbeing, life satisfaction, increased healthcare utilization, and decreased work productivity (Loebach Wetherell et al., 2004; Stein & Heimberg, 2004; Wittchen, 2002).

Both Cognitive Behavioral Therapy (CBT) and pharmacotherapy are 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,

☆ Clinical Trial Registry Number: ISRCTN89276818 (http://www.isrctn.com/ISRCTN89276818).

* Corresponding author. Big Health, Runway East London Bridge, 20 St. Thomas St., London SE1 9RS, UK.
E-mail address: chris.miller@bighealth.com (C.B. Miller).

https://doi.org/10.1016/j.jbtep.2020.101609
Received 23 December 2019; Received in revised form 10 July 2020; Accepted 17 August 2020
Available online 8 September 2020

2011). CBT has larger overall treatment effects ($g = 0.76$) than medication ($g = 0.38$; Carl et al., 2020), fewer side effects (Baldwin et al., 2014), is more acceptable (Deacon & Abramowitz, 2005), and better tolerated (Mitte, 2005) by patients with GAD. Typical CBT components for GAD include cognitive restructuring, imaginal exposures, behavioral experiments, stimulus control, applied relaxation, self-monitoring, and psychoeducation (Craske & Barlow, 2006; Shafran, Brosan, & Cooper, 2013); some protocols also include mindfulness and/or acceptance-based techniques (e.g., Roemer, Orsillo, & Salters-Pedneault, 2008). Despite the favorable benefit/harm profile of CBT for GAD, accessibility of CBT is limited by insufficient numbers of trained therapists, costs, waiting lists, scheduling, distance from services, and perceived stigma (Comer & Barlow, 2014; Gunter & Whittal, 2010).

'*Digital CBT*' may help overcome such barriers to access because, computers, tablets, and smartphones are now pervasive in society (Holmes et al., 2018). Smartphones are owned by 81% of Americans (Pew Research Center, 2019) and are a promising 'next-generation' platform for delivering psychological therapy (Bhugra et al., 2017). Digital CBT programs for GAD have predominantly been web-based (i.e., optimized for use on a computer with internet access; e.g., Robinson et al., 2010; Carlbring et al., 2011) and 'guided' (i.e., requiring support from trained technicians or clinicians; Richards et al., 2016; Titov et al., 2009; Titov, Andrews, Johnston, Robinson, & Spence, 2010), limiting accessibility. Existing programs tend to lack advanced design and functionality possible with smartphones, such as automated yet personalized real-time support, which may be conducive to engagement and skill acquisition (Linardon, Cuijpers, Carlbring, Messer, & Fuller-Tyszkiewicz, 2019). Even with such limitations, web-based CBT interventions have shown large effect size improvements for GAD symptoms ($d = 0.91$; Richards, Richardson, Timulak, & McElvaney, 2015). Fully automated, smartphone-based CBT for GAD has the potential to increase access to care as well as enhance engagement and skill acquisition, possibly improving clinical effects.

This study examined feasibility and preliminary efficacy of *Daylight*™, a novel, fully automated, digital CBT intervention to treat symptoms of GAD. *Daylight* is delivered through a smartphone application ('app') and personalized based on users' responses. Prior to this study, initial pilot data collection with 18 participants who accessed *Daylight* using an earlier version (release 2018 for iOS), identified intervention and study issues which were resolved prior to this study (see supplemental material).

Primary objectives were to examine feasibility of *Daylight* in terms of retention, safety, adherence, and acceptability; and preliminary efficacy to reduce symptoms of GAD. Secondary objectives were to examine effects on worry, depressive symptoms, sleep, wellbeing, participant-specific quality of life, work functioning, and CBT skill acquisition. These objectives adhered to good practice guidelines for preliminary evaluations of a novel intervention (e.g., Stage Model of Behavioral Therapies; Rounsaville, Carroll, & Onken, 2006).

2. Materials and methods

2.1. Participants

Inclusion criteria were: 1) adults (aged 18+yrs. mean age 43yrs. range 19–65yrs.), 2) scoring ≥10 on the 7-item Generalized Anxiety Disorder questionnaire (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006), indicating moderate-to-severe symptoms, 3) screening positive for probable GAD diagnosis using a digital version of the Mini International Neuropsychiatric Interview (MINI; Sheehan, 2014) version-7 for the DSM-5, and 4) either not on prescription medication for anxiety, depressive symptoms, or sleep, or on a stable dose for at least 4-weeks. Exclusion criteria were: 1) past or present psychosis, schizophrenia, bipolar disorder, seizure disorder, or substance use disorder, 2) trauma to the head or brain damage, 3) severe cognitive impairment, 4) serious physical health concerns necessitating surgery or with prognosis <6 months, and 5) pregnancy. A total of 21 adults (20f) were recruited. The sample size was not statistically calculated and determined from previous literature on single-case experimental design studies (Franklin, Allison, & Gorman, 2014).

2.2. Design

The study comprised a randomized, multiple-baseline single-case experimental design (SCED). SCED is recommended for preliminary evaluation of novel psychological interventions as it allows for examination of treatment-related changes both within and between participants in small samples, whilst permitting experimental control (Barlow, Nock, & Hersen, 2009). The intervention was introduced following a randomly allocated baseline period of 2-, 4-, or 6-weeks duration. Primary anxiety outcomes were assessed daily and weekly to permit detection of fine-grained symptom changes; greater change being expected after introduction of the intervention (Barlow et al., 2009). Secondary outcomes included weekly measures of depressive symptoms and sleep. 'Global' secondary outcomes including the GAD-7 were assessed within-subjects; with questionnaires administered immediately before randomization at initial assessment, at post-intervention (6-weeks from intervention start), and at final follow-up (10-weeks from intervention start). Our approach adhered to guidelines including '*What Works Clearinghouse*' for single-case designs (Kratochwill et al., 2010), Single-Case Reporting Guideline In Behavioral Interventions (Tate et al., 2016), CONSORT extension for reporting N-of-1 trials: CENT (Shamseer et al., 2015), and reporting standards for quantitative research in psychology (Appelbaum et al., 2018).

2.3. Ethics

The University of Oxford Medical Sciences Inter-Divisional Ethics Committee (reference R58113/RE001) gave ethical approval and the study was prospectively registered with the ISRCTN registry (#89276818, https://doi.org/10.1186/ISRCTN89276818). The full trial protocol is available on request to the corresponding author.

2.4. Procedure

Advertisements were placed on social media in the UK, directing interested participants to an online information sheet and stipulated individuals required access to a smartphone. Individuals consented to and completed a short survey for eligibility. Eligible participants were telephone-screened by a postgraduate-level research psychologist, trained and supervised by three clinical psychologists (MLD, RS, & JRC). The researcher explained study procedures, verified eligibility and probable GAD diagnosis from study entry measures, including the MINI, and answered questions. Participants completed an online consent form and initial questionnaire assessments. Participants were then randomized automatically and independent of the study team by Qualtrics (Qualtrics, 2019) to one of three baseline periods (2-, 4-, or 6-weeks). To account for greater attrition in the longer baseline groups, we weighted randomization by blocks of 5, 7, and 9 to the 2-, 4-, and 6-week groups, respectively.

All subsequent aspects of the study were completed online, and assessments were captured using Qualtrics. Treatment process measurement required participants to complete a daily question about their anxiety captured in response to a text message sent at 2p.m. during baseline, and for 6-weeks during the intervention phase. Weekly, online surveys measured anxiety, depressive symptoms and sleep during baseline, intervention, and for 4-weeks during the follow-up phase. The final post-intervention survey, and final follow-up survey included global measures administered at the initial assessment. All participants immediately received the intervention after completing their baseline phase, and access was not withdrawn.

The study lasted between 12 and 16 weeks, depending on baseline allocation. Participants were compensated in Amazon vouchers and payment was in line with Oxford University ethical review requirements where participants are paid for their time. Participants received £10 ($12.30 USD) in vouchers for each week they completed assessments and £15 ($18.50) for each global assessment (initial, post-intervention and follow-up). The total was dependent on their baseline allocation (2-, 4-, or 6-weeks) and ranged from £165 ($203) to £205 ($252).

2.5. Digital CBT intervention

The intervention, *Daylight* (www.bighealth.com/daylight: iOS and Android release 2019) is a fully automated CBT-based program developed specifically for smartphone delivery. *Daylight* is a voice-led experience in which a virtual therapist guides the user through CBT skills while providing empathic support and personalized feedback. *Daylight* was developed by experts in anxiety disorders in conjunction with designers, filmmakers, podcast producers, and animators to create an efficacious and engaging program. The design and features include animated psychoeducational videos, personalized in-the-moment support and technique recommendations, and brief practice exercises. *Daylight* consists of four core modules (approximately 10–20 min in length) and includes stimulus control, applied relaxation, cognitive restructuring, and imaginal exposure (see supplemental table 1). Tailored feedback and personalized troubleshooting are provided based on user input. Modules are accessed sequentially and can be repeated. Users can complete shorter (approximately 5 min) practice exercises for each module. Brief in-app assessments quantify levels of worry, anxiety, mood and sleep, and give weekly progress feedback. The program encourages and reinforces daily use (e.g., practise a technique) as well as real world implementation of therapeutic content (i.e., use of techniques outside of the app). Though not reported here, implementation is captured each time the app is used by asking if users applied techniques on their own.

2.6. Measures

2.6.1. Feasibility

Feasibility was assessed by measurement of 1) retention, 2) adherence, 3) satisfaction, 4) safety, and 5) credibility. For retention and adherence, we used objective engagement statistics from *Daylight* to determine those who completed all four modules (retention), and the total number of modules completed (adherence). Satisfaction was assessed at post-intervention by qualitative and quantitative questions (see supplemental material). Safety included the occurrence of any adverse events reported throughout the study period, spontaneously or in response to open-ended questions by participants to the study coordinator during any correspondence from consent until the final assessment. Clinical Psychologists were available to follow-up reports of adverse events directly with participants by telephone. Safety was also assessed at post-intervention, by asking participants to rate the occurrence of potential unwanted symptoms related to the intervention (e.g., low mood, feeling agitated, headache, fatigue) using a Modified Symptom Checklist (Kyle, Morgan, Spiegelhalder, & Espie, 2011). Credibility was evaluated by the Credibility/Expectancy Questionnaire (Devilly & Borkovec, 2000), assessed in the first weekly survey during the intervention period.

2.6.2. Preliminary efficacy

Daily levels were examined using a single-item question, based on Loerinc's (2018) daily measure of: '*On average over the past 24 h, how anxious or fearful have you felt?*'. Responses were captured using a digital visual analogue sliding scale ranging from 0 (not at all) to 10 (extremely). Measures of anxiety (GAD-7; Spitzer et al., 2006), depressive symptoms (Patient Health Questionnaire 9-item: PHQ-9; Kroenke, Spitzer, & Williams, 2001), and sleep (7-item version of the Sleep Condition Indicator: SCI; Espie et al., 2014) were assessed weekly during baseline, intervention, and follow-up periods. The following measures were administered at all global assessment batteries (initial assessment, post-intervention, and final follow-up) and assessed worry, wellbeing, participant-specific quality of life, work productivity, and CBT skill acquisition: Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990), Warwick Edinburgh Mental Wellbeing Scale (WEMWBS; Tennant et al., 2007), Patient-Generated Index (Ruta, Garratt, Leng, Russell, & MacDonald, 1994), Work Productivity and Activity Impairment index (WPAI; Reilly, Zbrozek, & Dukes, 1993), and the CBT Skills Questionnaire (CBTSQ; Jacob, Christopher, & Neuhaus, 2011). See supplemental material for more details and supplemental table 2 for an overview.

2.7. Data analysis

2.7.1. Feasibility

Descriptive statistics were used to report intervention retention, adherence, safety, credibility, and satisfaction. Qualitative feedback assessing intervention satisfaction were analyzed using thematic analysis to examine themes across participants' responses. This approach followed steps outlined by Braun and Clarke (2006).

2.7.2. Preliminary efficacy to improve daily and weekly anxiety, depressive symptoms, and sleep outcomes

Analyses were conducted in accordance with established guidelines for SCED research and used a combination of statistical methods and visual inspection (Barlow et al., 2009; Tate et al., 2016). Data were analyzed as a series of single-case statistical analyses and then combined into multilevel analyses to establish overall patterns of baseline and trend changes and to assess individual differences. Multilevel analyses were conducted with the lmer procedure in the R statistical package, *lme4* (Bates, Mächler, Bolker, & Walker, 2015). To quantify between-participant variability, we report the intraclass correlation coefficient (ICC) for mixed models. This use of ICC is different from the use of an ICC in inter-rater reliability assessment because it evaluates the between-participant variance for different slopes and intercepts. There are no quantified cut-offs for this use of ICC for interpretation, instead, higher scores between 0 and 1 indicate increased between-participant variance (Franklin et al., 2014). Each individual's data were examined as a single-case by means of time series analysis, control and sequence charts (Eubanks-Carter, Gorman, & Muran, 2012), post-hoc tests (Holm's correction) of baseline, intervention, and follow-up (weekly only), phase means, and associated d and r effect size measures. For visual inspection, outcomes were graphed for each participant and the effect of the intervention was examined by visually comparing the magnitude (i.e., level) and rate of change (i.e., slope) during intervention and follow-up phases compared to the baseline phase. Visual analyses were implemented with graphics in the R statistical package *SSD for R* (R Core Team, 2018).

2.7.3. Efficacy to improve global outcomes

Changes in anxiety associated with GAD, sleep, depressive symptoms, worry, wellbeing, participant-specific quality of life, work performance, and CBT skill acquisition across global timepoints (initial assessment, post-intervention, and final follow-up) were assessed using repeated-measures ANOVAs. These were followed-up with paired-samples t-tests, Cohen's d effect sizes, and 95%CIs to examine initial assessment to post-intervention and initial assessment to final follow-up within-subject change. Ancillary analyses examined rates of clinically significant change from initial assessment to both post-intervention and final follow-up for anxiety (GAD-7), depressive symptoms (PHQ-9), and sleep (SCI-8) outcomes for each participant (for those who experienced clinically significant symptoms at baseline only). Participants did not have clinically significant anxiety, depressive symptoms or sleep difficulty if they scored <10 on the GAD-7 (Spitzer et al., 2006), <10 on the

PHQ-9 (National Collaborating Centre for Mental Health, 2018), and >16 on the SCI-8 (Espie et al., 2014), respectively. Clinically significant change was considered reliable if participants scored below the above thresholds on a specific measure (as defined) and demonstrated a change score greater than the known unreliability of the measure (Richards & Borglin, 2011) [for GAD-7 this involved reductions of ≥5 (Spitzer et al., 2006), for PHQ-9 reductions of ≥6 (National Collaborating Centre for Mental Health, 2018), and for the SCI-8 increases of ≥7 (Espie et al., 2018)].

3. Results

3.1. Participant flow

Participant flow by baseline duration and individual participant characteristics are presented in Fig. 1 and Table 1, respectively. In accordance with CONSORT, overall participant flow is reported using the CONSORT CENT diagram. In total, 43 participants were assessed online for eligibility against study inclusion/exclusion criteria. Thirty-two were eligible and invited to the telephone screen. Of these, 21 consented into the study, and completed the initial questionnaire assessment. Participants were randomized to baseline periods of 2- (n = 5), 4- (n = 6), or 6-weeks (n = 10) in duration. All participants randomized, downloaded, and started *Daylight* (13 accessed the app using Android and 8 accessed using iOS devices) and no participants withdrew. Recruitment started on September 3, 2018, was completed on April 30, 2019, and the last follow-up occurred on July 12, 2019.

3.2. Feasibility

3.2.1. Retention and adherence

Sixteen participants (76%) completed all four modules of the intervention, 17 (81%) completed three modules, 18 (86%) completed two modules, and all 21 (100%) completed at least one. On average, participants completed 3.43 (SD = 1.12) out of 4 modules and the median was 4. The mean number of pieces of therapeutic content completed in the app was 13 (SD = 11; range of 1–37). Content includes the number of times modules (which could be repeated) and practice exercises were completed.

3.2.2. Safety

Throughout the study period, no serious adverse events or adverse events were reported. In terms of occurrence of unwanted symptoms related to the intervention, assessed using a Modified Symptom Checklist, nine of the 21 participants (43%) endorsed unwanted symptoms during the intervention including: agitation, low mood, fatigue/exhaustion, and reduced motivation and/or energy (supplemental table 3).

3.2.3. Credibility

Impressions of intervention credibility were positive with moderate-high scores (M = 19.30, SD = 4.78, range = 10–26, n = 20).

3.2.4. Satisfaction

Mean scores indicated participants were mostly satisfied with the intervention (M = 6.20, SD = 1.99, range = 2–10, n = 20). One participant (5%) was completely satisfied, most (n = 7) responded with a 7, indicating satisfaction, and none were totally dissatisfied. From qualitative reports, participants found *Daylight* enjoyable and effective, describing the program as helpful to develop coping strategies to challenge and address anxious thoughts and worries. They felt *Daylight* integrated easily into their daily life. Barriers to intervention use were highlighted, including time constraints, forgetfulness and frustration due to technical problems. Suggested further changes included addressing technical issues, module reminders, and a silent mode. Many indicated that *Daylight* addressed their needs, however some felt it was not specific enough for their symptoms of anxiety.

3.3. Preliminary efficacy

3.3.1. Efficacy to improve daily anxiety and worry

A total of 1491 daily anxiety ratings were provided by 21 participants, representing a 97% completion rate (out of 1540 possible assessments). The number of daily reports per participant ranged from 46 to 88 days. Across participants, using multilevel modelling, there was a significant effect of time on daily anxiety scores, with overall scores decreasing from baseline (M = 6.46, SE = 0.30) to intervention (M = 5.45, SE = 0.31), (t (19.96) = 4.15, p < .001), see Fig. 2. The ICC among individual's anxiety ratings was 0.37, indicating some between-participant variability. There was no significant change during baseline (p = .185). Of the 21 participants, 11 showed significant overall phase improvement in daily anxiety levels from baseline to intervention phase and one showed significant deterioration in daily anxiety levels from baseline to intervention phase. Visual inspection of individual participant sequence charts (Fig. 3) corroborated these findings.

3.3.2. Efficacy for weekly anxiety, depressive symptoms, and sleep outcomes

3.3.2.1. Weekly anxiety symptoms (GAD-7). A multilevel analysis, modelling study employed phase as a fixed effect and individual intercepts and slopes as random effects, indicated a statistically significant relationship between GAD-7 levels and study phase (baseline, intervention, and follow-up), (F (2,14.36) = 5.41, p = .018, r = 0.28). The ICC was 0.66, indicating considerable between-participant variability. There was a significant effect of time on weekly anxiety scores, with overall scores decreasing from baseline (M = 12.57, SE = 0.97) to intervention phase (M = 10.25, SE = 0.78), t (19.52) = 2.78, p = .024, and follow-up (M = 9.11, SE = 1.17), t (16.88) = 3.15, p = .017. Further analysis of GAD ratings at each week of the baseline phase was small and statistically non-significant (GAD = 12.38 + 0.05*Week; (F (1,59) = 0.07, p = .786; r = 0.00). The ICC for baseline GAD-7 ratings was 0.8, indicating large between-participant variability. Of the 21 participants, 11 showed significant phase improvement in weekly anxiety levels from baseline to intervention phase, 16 showed significant phase improvement in weekly anxiety levels from baseline to follow-up, and six showed significant phase improvement in weekly anxiety levels from intervention to follow-up phase.

3.3.2.2. Weekly depressive symptoms (PHQ-9). A similar multilevel analysis indicated a statistically significant relationship between PHQ-9 levels and phase (F (2,19.73) = 20.6, p < .001, r = 0.44), the ICC was 0.64, indicating considerable between-participant variability. There was a significant effect of time on weekly depressive symptom scores, decreasing from baseline (M = 14.41, SE = 0.96) to intervention phase (M = 9.76, SE = 0.87), t (20.51) = 5.75, p < .001, and baseline to follow-up (M = 8.32, SE = 1.13), t (19.76) = 5.92, p < .001. There was no statistically significant trend for PHQ-9 scores during the baseline phase (PHQ-9 = 14.28–.32*week; (F (1,16.68) = 2.27, p = .151, r = 0.10). The ICC was 0.73, indicating considerable between-participant differences. Of the 21 participants, 12 showed significant phase improvement in weekly depressive symptoms from baseline to intervention, 15 showed significant phase improvement in weekly depressive symptoms from baseline to follow-up, six showed significant phase improvement from intervention to follow-up, and two showed significant phase deterioration from intervention to follow-up.

3.3.2.3. Weekly sleep (SCI-7). The multilevel analysis, indicated a statistically significant relationship between levels and phase (F (2,18.98) = 15.77, p < .001, r = 0.30, ICC = 0.80). There was a significant effect of time on weekly sleep scores, with overall scores improving from baseline

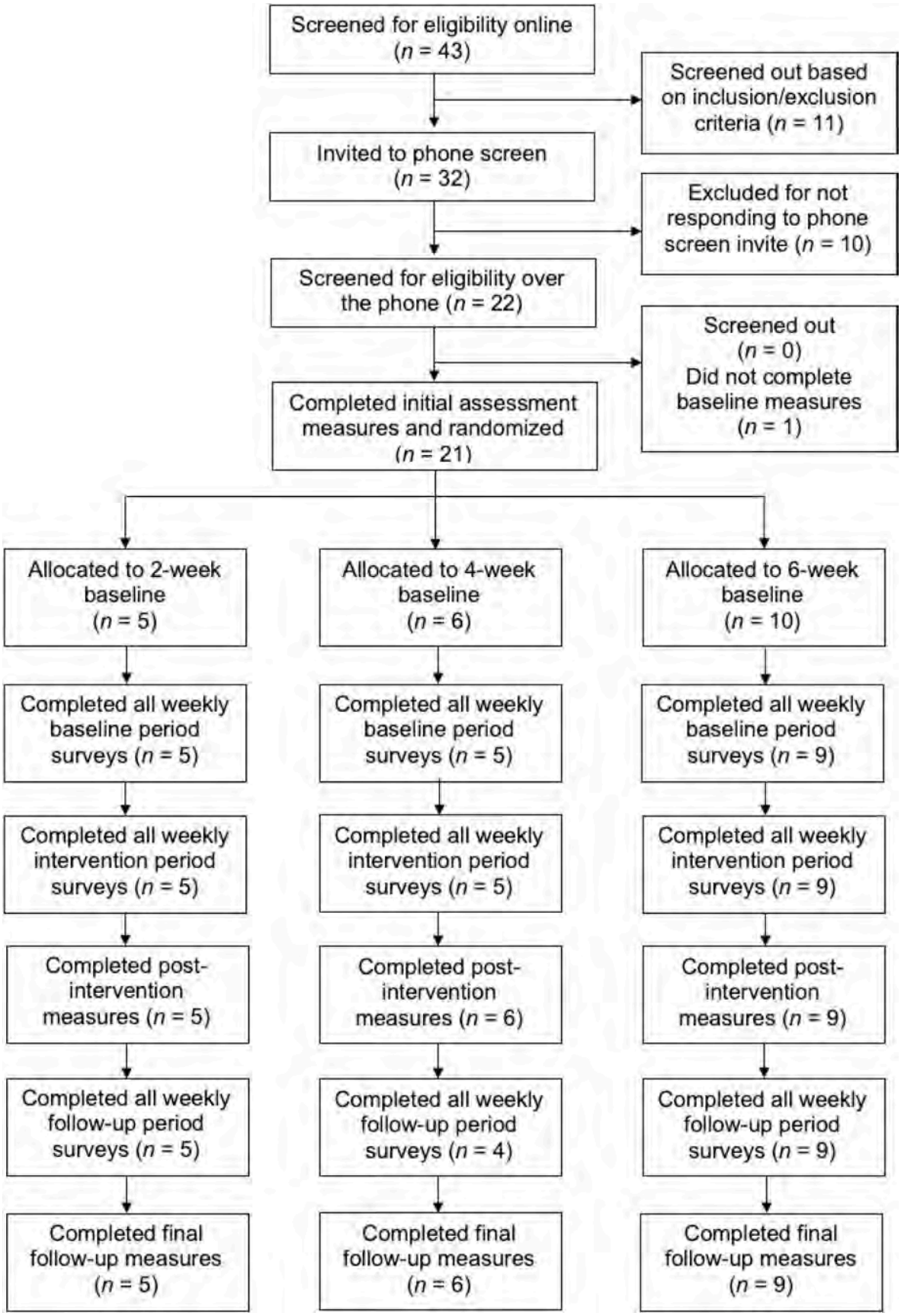

Fig. 1. Participant flow through the study.

Table 1
Participant characteristics at baseline, randomized baseline duration, and intervention credibility.

	Age (years)	Gender	Ethnicity	Educational attainment	Employment status	Duration of anxiety difficulty	No. days of medication in last week	No. treatment provider visits in last week	Randomized baseline duration	Credibility
P1	38	Female	White	Undergraduate/ bachelor's degree	Part-time employed	>3 years	0	1	4-weeks	11
P2	39	Female	White	Postgraduate or professional degree	Unemployed	>3 years	0	0	6-weeks	.
P3	48	Female	White	College	Unemployed	>3 years	0	0	4-weeks	26
P4	56	Female	White	Secondary school/ high school graduate	Unemployed	>3 years	7	0	2-weeks	10
P5	30	Male	White	Undergraduate/ bachelor's degree	Unemployed	>3 years	2	0	2-weeks	25
P6	49	Female	White	Undergraduate/ bachelor's degree	Full-time employed	>3 years	7	1	6-weeks	23
P7	61	Female	White	Secondary school/ high school graduate	Unemployed	>3 years	0	0	2-weeks	23
P8	65	Female	White	Postgraduate or professional degree	Retired	>3 years	7	1	4-weeks	19
P9	21	Female	White	College	Full-time employed	1–3 years	4	0	6-weeks	19
P10	39	Female	White	Postgraduate or professional degree	Part-time employed	>3 years	0	0	6-weeks	21
P11	62	Female	White	Postgraduate or professional degree	Full-time employed	>3 years	0	0	6-weeks	26
P12	27	Female	White	College	Part-time employed	>3 years	7	0	6-weeks	16
P13	26	Female	White	Secondary school/ high school graduate	Full-time homemaker or carer	>3 years	6	0	4-weeks	20
P14	43	Female	White	Postgraduate or professional degree	Part-time employed	>3 years	0	0	6-weeks	23
P15	57	Female	White	College	Part-time employed	>3 years	0	0	6-weeks	13
P16	31	Female	White	Postgraduate or professional degree	Full-time employed	>3 years	0	0	2-weeks	13
P17	63	Female	White	No formal qualifications	Part-time employed	>3 years	7	0	4-weeks	22
P18	50	Female	White	Secondary school/ high school graduate	Full-time homemaker or carer	>3 years	7	1	4-weeks	19
P19	28	Female	White	Postgraduate or professional degree	Full-time employed	>3 years	7	0	6-weeks	16
P20	19	Female	White	College	Full-time student	>3 years	7	0	6-weeks	20
P21	52	Female	White	College	Unemployed	>3 years	0	0	2-weeks	21

Note: Credibility was assessed in the first weekly survey during the intervention period through the Credibility/Expectancy Questionnaire.

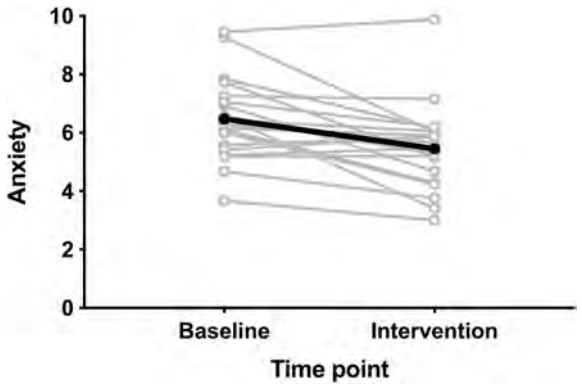

Fig. 2. Daily symptoms of anxiety over time.
Fig. 2 Note: Average (dark line) and individual (lighter lines) anxiety and worry symptoms at baseline (2-, 4-, or 6-weeks duration) and intervention (6-weeks) periods. Symptoms were measured by a single item: *'On average over the past 24 h, h'*. Responses were captured using a digital visual analogue scale ranging from 0 (not at all) to 10 (extremely).

($M = 10.50$, $SE = 1.26$) to intervention phase ($M = 14.99$, $SE = 1.51$), t (19.42) = 4.74, $p < .001$, and baseline to follow-up ($M = 15.41$, $SE = 1.66$), t (19.71) = 4.98, $p < .001$. There was no statistically significant baseline trend (SCI-7 = 9.38–.32*week; (F (1,27.39) = 3.14, $p = .087$, $r = 0.10$, ICC = 0.69). Of the 21 participants, nine showed significant phase improvement in weekly sleep levels from baseline to intervention phase, 12 showed significant phase improvement in weekly sleep levels from baseline to follow-up, four showed significant phase improvement from intervention to follow-up, and one showed significant phase deterioration from intervention to follow-up. In each analysis, there was no evidence of baseline trends but there were significant between-phase effects.

3.3.2.4. Visual inspection. Graphs displaying each participant's average baseline, intervention, and follow-up phase data for weekly anxiety (Fig. 4a), depressive symptoms (Fig. 4b), and sleep (Fig. 4c) illustrates these trends and are consistent with the statistical findings.

3.3.3. Efficacy to improve global outcomes

Table 2 presents means, SDs, repeated-measures ANOVAs with accompanying Cohen's d, for secondary outcomes at global timepoints (initial assessment, post-intervention and final follow-up). There was a significant and large reduction in anxiety (Fig. 5), depressive symptoms

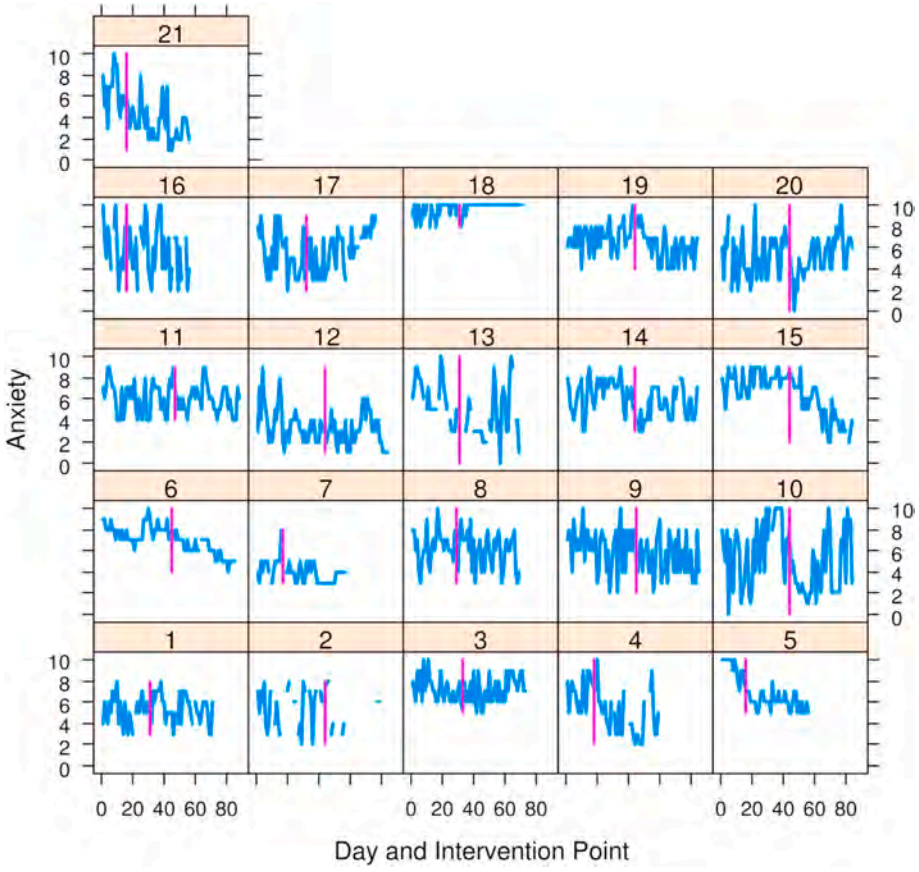

Fig. 3. Individual participant sequence charts of daily symptoms of anxiety over time.
Fig. 3 Note: Anxiety symptoms were assessed daily at baseline (2-, 4-, or 6-weeks duration) and during intervention (6-weeks) periods. The vertical reference line marks the start of the intervention. Symptoms of anxiety were measured by a single item: *'On average over the past 24 h, h'*. Responses were captured using a digital visual analogue scale ranging from 0 (not at all) to 10 (extremely).

and sleep difficulty over time (all $p < .001$). We also observed significant improvements over time for worry (PSWQ), wellbeing (WEMWBS) and CBT skills acquisition (CBTSQ) (all $p < .001$). For the Patient-Generated Index, there was a significant improvement in all three participant-specific areas over time ($p < .001$-.034). There were no significant improvements for work productivity outcomes of absenteeism and presenteeism (WPAI).

3.3.3.1. Clinically significant change. For anxiety (GAD-7), 14/20 (70%) no longer had clinically significant symptoms at post-intervention and 13/20 (65%) had both a clinically significant and reliable change. At final follow-up, this increased to 17/20 (85%) for clinical change and 16/20 (80%) for clinical and reliable change. Similarly, for those with depressive symptoms (PHQ-9) at baseline, 11/18 (61%) had both clinical and reliable change at post-intervention, and 8/18 (44%) maintained both at final follow-up. For those with sleep difficulty (SCI-8) at baseline, 8/20 (40%), had clinical change suggestive of healthy sleep at post-intervention and 7/20 (35%) had clinical and reliable change. At final follow-up, this increased to 9/20 (45%) for clinical change and 8/20 (40%) for clinical and reliable change. None showed reliable deterioration from initial assessment to post-intervention/follow-up for any of the outcomes above.

4. Discussion

This study used a randomized, multiple-baseline SCED to examine the feasibility and preliminary efficacy of a novel smartphone-based fully automated digital CBT therapeutic (*Daylight*) for symptoms of GAD. Results suggest *Daylight*, is feasible in terms of acceptability (uptake), engagement (retention and adherence), credibility, satisfaction, and safety. All participants downloaded and accessed therapeutic content, and 76% (16/21) completed all modules of the intervention. Overall impressions of credibility and satisfaction were positive. Qualitatively, participants said the intervention was integrated easily into their daily life. They suggested helpful improvements to technical functioning, module reminders, and increased program personalization. For safety, no adverse events were reported by participants at any time during this study. Participants did report some unwanted symptoms, including agitation, low mood, fatigue/exhaustion, and reduced motivation and/or energy. These reports, when measured in studies of psychological treatments, are in line with typical side-effects of in-person CBT, and other digital CBT interventions (Espie et al., 2019; Gullickson, Hadjistavropoulos, Dear, & Titov, 2019). CBT techniques involve intentionally confronting distressing thoughts and engaging in potentially uncomfortable new behaviors, therefore, experiencing temporary distress is common (Foa, Zoellner, Feeny, Hembree, & Alvarez-Conrad, 2002; Schermuly-Haupt, Linden, & Rush, 2018). Learning to face and respond more flexibly to distressing thoughts, physiological sensations, and emotions (e.g., through thought monitoring, exposure exercises, behavioral activation, etc.) is a primary proposed mechanism of CBT (Mennin, Ellard, Fresco, & Gross, 2013). In future work, a between-group comparison may determine if unwanted symptoms occur at a higher rate for those assigned to the intervention compared with a control condition.

For the daily and weekly anxiety outcomes, results provide preliminary support for the efficacy of *Daylight* to improve symptoms of GAD. Fine-grained daily assessments reduced significantly with the introduction of the intervention and not during baseline. Turning to weekly measures, we observed a significant improvement in weekly measures of anxiety (GAD-7). Visual inspection of individual participants' daily and weekly data verified these findings as improvements occurred specifically following the introduction of the intervention. For secondary outcomes, we found improvements to weekly measures of depressive symptoms (PHQ-9), and sleep difficulty (SCI-7). These

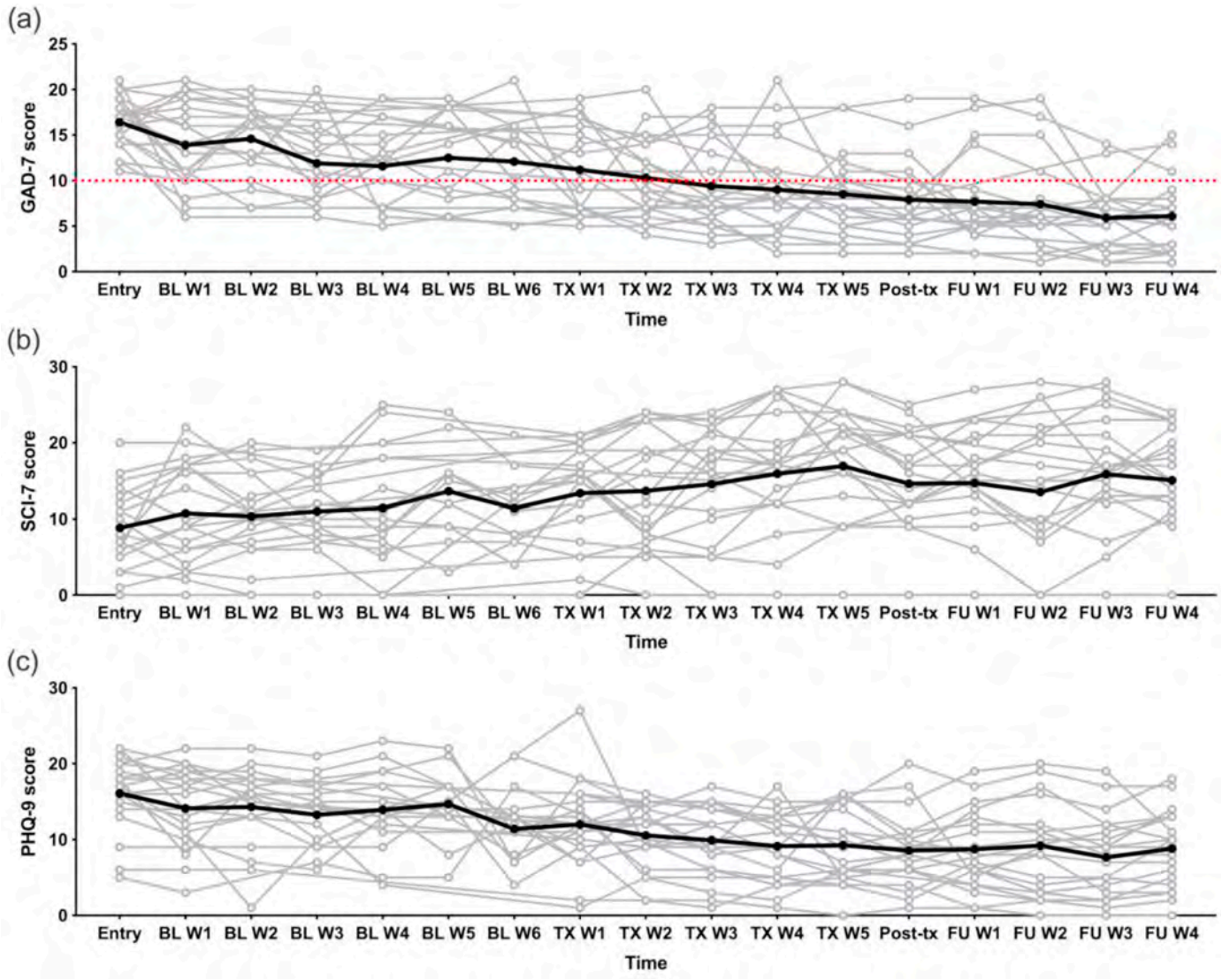

Fig. 4. Weekly average (dark line) and individual (lighter lines) symptoms of anxiety (a), sleep difficulty (b), and depressive symptoms (c) over time.
Fig. 4 Note: Average (dark line) and individual (lighter lines) for all 21 participants. Symptoms of Anxiety measured by the 7-item Generalized anxiety disorder questionnaire (GAD-7), depressive symptoms by the 9-item Patient health questionnaire (PHQ-9), and sleep by the 7-item Sleep condition indicator (SCI-7): greater scores indicate better sleep. Outcomes were assessed weekly at baseline (2-, 4-, or 6-weeks duration), during intervention (6-weeks), and follow-up (4-weeks) study periods. All questionnaires were modified to examine symptoms over the previous week. The dotted red line in Fig. 3 (a) illustrates the clinical cut-off for anxiety (score of 10 or higher).

findings were further supported at the level of individual participants, as the majority significantly improved at follow-up for all three outcomes (GAD-7, PHQ-9, and SCI-7). Overall initial assessment to post-intervention (6-weeks from intervention start) and final follow-up (10-weeks from intervention start) reductions on global anxiety symptoms (GAD-7) were large and clinically meaningful with 65% (13/20) of participants meeting criteria for clinically significant and reliable change (Richards & Borglin, 2011). This increased to 80% (16/20) at final follow-up. The greatest number of participants demonstrated a clinically significant change with the GAD-7 compared with secondary measures of depressive symptoms (PHQ-9) and sleep difficulty (SCI-8). This is in line with our primary hypothesis, that the intervention aims to target anxiety associated with GAD.

The large effects observed from the global GAD-7 assessment compared with the smaller effects observed in the daily-item measure may be because the daily measure was composed of a single item and captures more general daily stress to a greater degree than specific symptoms of GAD. In turn, this may make it more difficult to show reliable change. Nevertheless, the daily effects were directionally consistent with the weekly and global findings. We observed significant overall within-subject improvements across participants for measures of worry (PSWQ), wellbeing (WEMWBS), and CBT skills acquisition (CBTSQ). All improvements were first detected at post-intervention and maintained at follow-up. It is encouraging to observe improvements in worry, as it is the defining feature of GAD, and previous research suggests worry is a primary maintaining mechanism of the disorder (Broschot, Van Dijk, & Thayer, 2007; Llera & Newman, 2010; Stapinski, Abbott, & Rapee, 2010). The increase in CBT skills further suggests individuals acquire new skills as a result of digital CBT that support the improvements in symptoms. The development of CBT skills is critical to individuals' obtaining lasting effects, a key benefit of CBT. Longer duration follow-up observations are now needed. We also observed significant improvements (at both post-intervention and final follow-up) in ratings of all three top areas of life affected by anxiety identified by participants using the Patient-Generated Index. These effects are noteworthy because they suggest *Daylight* has potential to improve

Table 2
Means, standard deviations, *t*-values, and repeated measures effect size scores for all global outcomes assessed at initial assessment before baseline period randomization, 6-weeks post-intervention, and at 10-weeks final follow-up from intervention start.

Measure	Mean (SD); *n*			*F*-value (df)	*t*-value (df); Cohen's *d* (95% CI)	
	Initial	Post	Follow-up	Initial-Post-Follow-up	Initial-Post	Initial-Follow-up
GAD-7	16.55 (2.80); *n* = 20	7.85 (4.45); *n* = 20	6.10 (3.96); *n* = 20	F (2, 38) = 64.88, p < .001	t (19) = 8.63, p < .001; d = 2.29 (1.40, 3.19)	t (19) = 10.15, p < .001; d = 3.04 (1.91, 4.17)
PSWQ	72.45 (5.19); *n* = 20	61.75 (10.40); *n* = 20	60.30 (11.33); *n* = 20	F (2, 38) = 20.21, p < .001	t (19) = 5.65, p < .001; d = 1.15 (0.61, 1.69)	t (19) = 5.69, p < .001; d = 1.21 (0.64, 1.78)
PHQ-9	16.65 (4.02); *n* = 20	8.55 (4.78); *n* = 20	8.80 (5.11); *n* = 20	F (2, 38) = 39.58, p < .001	t (19) = 7.51, p < .001; d = 1.82 (1.07, 2.57)	t (19) = 7.19, p < .001; d = 1.71 (0.99, 2.42)
WEMWBS	31.40 (6.40); *n* = 20	39.55 (7.96); *n* = 20	42.55 (10.10); *n* = 20	F (2, 38) = 25.21, p < .001	t (19) = 5.77, p < .001; d = 1.11 (0.59, 1.62)	t (19) = 5.27, p < .001; d = 1.27 (0.65, 1.90)
CBTSQ	37.70 (10.72); *n* = 20	49.60 (11.71); *n* = 20	51.05 (12.88); *n* = 20	F (2, 38) = 22.10, p < .001	t (19) = 5.81, p < .001; d = 1.06 (0.57, 1.55)	t (19) = 4.94, p < .001; d = 1.12 (0.55, 1.69)
SCI-7	8.30 (4.74); *n* = 20	14.60 (7.02); *n* = 20	15.05 (6.96); *n* = 20	F (2, 38) = 23.52, p < .001	t (19) = 5.34, p < .001; d = 0.98 (0.51, 1.46)	t (19) = 5.34, p < .001; d = 1.08 (0.56, 1.61)
SCI-8	8.80 (5.10); *n* = 20	15.1 (7.19); *n* = 20	15.55 (7.05); *n* = 20	F (2, 38) = 23.52, p < .001	t (19) = 5.34, p < .001; d = 0.96 (0.49, 1.42)	t (19) = 5.34, p < .001; d = 1.05 (0.54, 1.57)
Patient-Generated index 1st area of life affected by anxiety	38.24 (25.55); *n* = 17	55.88 (24.25); *n* = 17	55.88 (26.93); *n* = 17	F (2, 32) = 4.63, p = .017	t (16) = 2.32, p = .034; d = 0.71 (0.06, 1.35)	t (17) = 2.48, p = .024; d = 0.70 (0.10, 1.29)
Patient-Generated index 2nd area of life affected by anxiety	35.88 (18.39); *n* = 17	55.29 (23.48); *n* = 17	54.71 (24.27); *n* = 17	F (2, 32) = 12.62, p < .001	t (16) = 4.10, p = .001; d = 0.90 (0.37, 1.43)	t (17) = 4.19, p = .001; d = 0.77 (0.33, 1.22)
Patient-Generated index 3rd area of life affected by anxiety	32.94 (20.24); *n* = 17	54.12 (21.81); *n* = 17	60.59 (23.84); *n* = 17	F (2, 32) = 13.72, p < .001	t (16) = 5.54, p < .001; d = 1.00 (0.51, 1.50)	t (17) = 4.42, p < .001; d = 1.28 (0.57, 2.00)
WPAI Absenteeism in past week	0.24 (0.33); *n* = 11	0.11 (0.30); *n* = 11	0.11 (0.30); *n* = 11	F (2, 20) = 1.89, p = .178	t (10) = 1.33, p = .214; d = 0.42 (−0.23, 1.07)	t (10) = 1.48, p = .169; d = 0.42 (−0.17, 1.01)
WPAI Presenteeism in past week	45.45 (25.83); *n* = 11	30.91 (26.25); *n* = 11	41.82 (33.71); *n* = 11	F (2, 20) = 1.21, p = .319	t (10) = 1.68, p = .124; d = 0.56 (−0.14, 1.26)	t (10) = 0.34, p = .742; d = 0.12 (−0.58, 0.82)

Note: CBTSQ = Cognitive Behavioral Therapy Skills Questionnaire; GAD-7 = Generalized Anxiety Disorder questionnaire; PHQ-9 = Patient Health Questionnaire; PSWQ = Penn State Worry Questionnaire; SCI-8 = Sleep Condition Indicator; SD = Standard Deviation; WEMWBS = Warwick Edinburgh Mental Wellbeing Scale; WPAI = Work Productivity and Activity Impairment index.

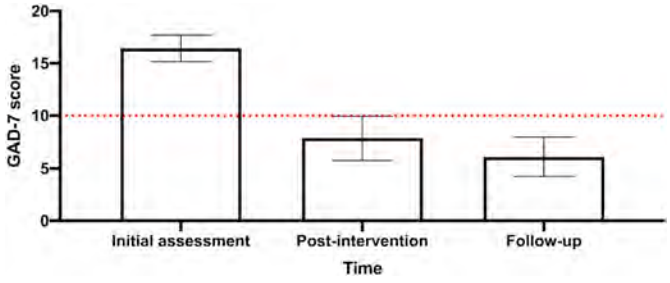

Fig. 5. Global anxiety symptoms over time.
Fig. 5 Note: Average symptoms of anxiety measured by the 7-item Generalized anxiety disorder questionnaire (GAD-7) assessed at initial assessment (n = 21), 6-weeks post-intervention (n = 20), and at 10-weeks final follow-up (n = 20) from intervention start. Error bars represent 95%CI. The dotted red line illustrates the clinical cut-off for anxiety (score of 10 or higher).

individual functioning beyond reducing symptoms. Prior research has shown that such improvements foster long-term resilience and may help reduce the risk of future depression (Grant, Guille, & Sen, 2013; Wood & Joseph, 2010). We did not detect any statistically significant improvements for WPAI-assessed measures of work absenteeism and presenteeism; however, less than one third of our study sample was employed full-time. The potential impact on workplace productivity requires further investigation in larger samples.

Overall, these preliminary findings support both the feasibility and early efficacy of *Daylight* in adults with moderate-to-severe GAD symptom severity. A digital therapeutic approach has the potential to provide full and immediate population level access to cognitive behavioral self-help for symptoms of GAD. *Daylight* will be made available as a medical benefit through health plans and large employers as a self-help approach and utilized alongside usual care. We used a rigorous experimental methodology evaluating symptoms at multiple assessment levels including global, weekly, and daily ecological assessments. The multiple-baseline design used in this study enables us to map improvements to symptoms based on the latency of change with intervention start. Further research is required and an RCT of *Daylight* would help test efficacy at the next level of rigour (Gu et al., 2020). Future studies with community-based clinical samples may also seek to understand in what way organizational, sociopolitical, and economic barriers faced by end-users (i.e., patients, providers, care systems) may be overcome to aid the implementation of digital therapeutics at scale (Graham, Lattie, & Mohr, 2019).

4.1. Limitations

A number of limitations should be acknowledged. Our sample was recruited online, rather than in-person. Further research is needed to understand if findings extend to other patient groups, including those

assessed in traditional healthcare settings and patients with different demographic characteristics. The present study did not specifically monitor and respond to elevated ratings on the PHQ-9. Providing such monitoring in future research could further support participant well-being. Our sample included 21 participants and only one participant identified as male, limiting the generalizability of the study findings. GAD is more common in females (Kessler et al., 2005), however, further research is required to establish whether *Daylight* is equally effective across sexes and further demographics. The study used text message prompts to remind participants for daily self-reports of their levels of anxiety. Although the intervention includes reminders and self-monitoring, it is possible that the study assessments had an additional impact on therapeutic outcomes. We did not specifically ask about previous use of psychotherapy for GAD but the study sample is potentially clinically-relevant because participants reported previous use of pharmacologic treatment for their anxiety. Although the experimental manipulation of baseline length across participants is designed to rule out other causes of symptom change than the study intervention, it is possible that other circumstances (e.g., access to other forms of psychotherapy) may have contributed to improvements.

5. Conclusions

A novel smartphone-based fully automated digital CBT intervention, *Daylight*, appears to be a feasible and safe therapy to help manage symptoms of anxiety and worry. Digital CBT was integrated into participants' daily lives and demonstrated preliminary efficacy to improve symptoms of GAD at daily, weekly and global outcome levels. The intervention was associated with improvements to secondary outcomes of worry, depressive symptoms, sleep difficulty, wellbeing, participant-specific measures of quality of life, and enabled acquisition of CBT skills. Further research is required to confirm the potential benefits of digital CBT for anxiety using RCT methodology.

Conflicts of interest

We wish to acknowledge any conflicts or potential conflicts of interest with the authors of this study. CAE is co-founder and Chief Medical Officer of Big Health Inc. and is a shareholder in the company. JRC and ALH are employed by Big Health Inc., are salaried by the company and are shareholders. CBM, RS, and MLD are employed by Big Health Inc. and are salaried by the company. AJH and JG were employed by Big Health Inc. BSG is a paid external statistical consultant of Big Health Inc. KHB, GMG and MGC report no conflicts of interest.

CRediT authorship contribution statement

Christopher B. Miller: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing - original draft, Writing - review & editing. **Jenny Gu:** Data curation, Formal analysis, Investigation, Methodology, Validation, Writing - original draft, Writing - review & editing. **Alasdair L. Henry:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing - original draft, Writing - review & editing. **Michelle L. Davis:** Methodology, Validation, Writing - original draft, Writing - review & editing. **Colin A. Espie:** Conceptualization, Data curation, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing. **Richard Stott:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Adrienne J. Heinz:** Conceptualization, Methodology, Project administration, Writing - original draft, Writing - review & editing. **Kate H. Bentley:** Methodology, Validation, Writing - original draft, Writing - review & editing. **Guy M. Goodwin:** Supervision, Project administration. **Bernard S. Gorman:** Formal analysis, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Michelle G. Craske:** Conceptualization, Methodology, Resources, Supervision, Writing - original draft, Writing - review & editing. **Jenna R. Carl:** Conceptualization, Formal analysis, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Declaration of competing interest

We would like to acknowledge the participants who took part in this research study. The intervention, *Daylight* was provided to all participants at no cost. The study was conducted at the University of Oxford, Nuffield Department of Clinical Neurosciences. The University of Oxford has a memorandum of understanding with Big Health for the conduct of joint research. This work was funded by Big Health Inc. The research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. Big Health Inc. was involved in the design and conduct of the study; collection, management, and interpretation of the data; preparation and review of the manuscript; and decision to submit the manuscript for publication. Other funders had no role in the design or conduct of the study, collection of data, data analysis, management, interpretation, or review or approval of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbtep.2020.101609.

References

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

Anxiety and Depression Association of America. (2020). Clinical practice review for GAD. Retrieved from https://adaa.org/resources-professionals/practice-guidelines-gad.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191.

Baldwin, D. S., Anderson, I. M., Nutt, D. J., Allgulander, C., Bandelow, B., den Boer, J. A., et al. (2014). Evidence-based pharmacological treatment of anxiety disorders, post-traumatic stress disorder and obsessive-compulsive disorder: A revision of the 2005 guidelines from the British association for psychopharmacology. *Journal of Psychopharmacology, 28*(5), 403–439.

Barlow, D., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston, MA: Allyn & Bacon.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Bhugra, D., Tasman, A., Pathare, S., Priebe, S., Smith, S., Torous, J., et al. (2017). The WPA- lancet psychiatry commission on the future of psychiatry. *Lancet Psychiatry, 4*(10), 775–818. https://doi.org/10.1016/S2215-0366(17)30333-4.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Brosschot, J. F., Van Dijk, E., & Thayer, J. F. (2007). Daily worry is related to low heart rate variability during waking and the subsequent nocturnal sleep period. *International Journal of Psychophysiology, 63*(1), 39–47. https://doi.org/10.1016/j.ijpsycho.2006.07.016.

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. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16933543.

Carl, E., Witcraft, S. M., Kauffman, B. Y., Gillespie, E. M., Becker, E. S., Cuijpers, P., et al. (2020). Psychological and pharmacological treatments for generalized anxiety disorder (GAD): A meta-analysis of randomized controlled trials. *Cognitive Behaviour Therapy, 49*(1), 1–21.

Carlbring, P., Maurin, L., Törngren, C., Linna, E., Eriksson, T., Sparthan, E., et al. (2011). Individually-tailored, internet-based treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy, 49*(1), 18–24. https://doi.org/10.1016/j.brat.2010.10.002.

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. *American Psychologist, 69*(1), 1–18. https://doi.org/10.1037/a0033582.

Craske, M. G., & Barlow, D. H. (2006). *Mastery of your anxiety and worry*. Oxford University Press.

Deacon, B. J., & Abramowitz, J. S. (2005). Patients' perceptions of pharmacological and cognitive-behavioral treatments for anxiety disorders. *Behavior Therapy, 36*(2), 139–145.
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.
Espie, C. A., Emsley, R., Kyle, S. D., Gordon, C., Drake, C. L., Siriwardena, A. N., et al. (2019). Effect of digital cognitive behavioral therapy for insomnia on health, psychological well-being, and sleep-related quality of life: A randomized clinical trial. *JAMA Psychiatry, 76*(1), 21–30.
Espie, C. A., Farias Machado, P., Carl, J. R., Kyle, S. D., Cape, J., Siriwardena, A. N., et al. (2018). The sleep condition indicator: Reference values derived from a sample of 200 000 adults. *Journal of Sleep Research, 27*(3), Article e12643.
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), Article e004183. https://doi.org/10.1136/bmjopen-2013-004183.
Eubanks-Carter, C., Gorman, B. S., & Muran, J. C. (2012). Quantitative naturalistic methods for detecting change points in psychotherapy research: An illustration with alliance ruptures. *Psychotherapy Research, 22*(6), 621–637. https://doi.org/10.1080/10503307.2012.693772.
Foa, E. B., Zoellner, L. A., Feeny, N. C., Hembree, E. A., & Alvarez-Conrad, J. (2002). Does imaginal exposure exacerbate PTSD symptoms? *Journal of Consulting and Clinical Psychology, 70*(4), 1022–1028. https://doi.org/10.1037//0022-006x.70.4.1022.
Franklin, R., Allison, D., & Gorman, B. (2014). *Design and analysis of single-case research.* Hoboken: Taylor and Francis.
Graham, A. K., Lattie, E. G., & Mohr, D. C. (2019). Experimental therapeutics for digital mental health. *JAMA Psychiatry, 76*(12), 1223–1224. https://doi.org/10.1001/jamapsychiatry.2019.2075.
Grant, F., Guille, C., & Sen, S. (2013). Well-being and the risk of depression under stress. *PloS One, 8*(7), Article e67395. https://doi.org/10.1371/journal.pone.0067395.
Gullickson, K. M., Hadjistavropoulos, H. D., Dear, B. F., & Titov, N. (2019). Negative effects associated with internet-delivered cognitive behaviour therapy: An analysis of client emails. *Internet Interventions.* , Article 100278. https://doi.org/10.1016/J.INVENT.2019.100278.
Gu, J., Miller, C. B., Henry, A. L., Espie, C. A., Davis, M. L., Stott, R., et al. (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.
Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., … Moulds, M. L. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *Lancet Psychiatry, 5*(3), 237–286.
Jacob, K. L., Christopher, M. S., & Neuhaus, E. C. (2011). Development and validation of the cognitive-behavioral therapy skills questionnaire. *Behavior Modification, 35*(6), 595–618. https://doi.org/10.1177/0145445511419254.
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. https://doi.org/10.1001/archpsyc.62.6.593.
Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., et al. (2010). Single-case designs technical documentation. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.
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. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11556941.
Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine, 146*(5), 317–325. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17339617.
Kyle, S. D., Morgan, K., Spiegelhalder, 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*(3), 325–336.
Llera, S. J., & Newman, M. G. (2010). Effects of worry on physiological and subjective reactivity to emotional stimuli in generalized anxiety disorder and nonanxious control participants. *Emotion, 10*(5), 640–650. https://doi.org/10.1037/a0019351.
Locke, A. B., Kirst, N., & Shultz, C. G. (2015). *Diagnosis and management of generalized anxiety disorder and panic disorder in adults* (Vol. 91). Retrieved from www.aafp.org/afpAmericanFamilyPhysician617.
Loebach Wetherell, J., Thorp, S. R., Patterson, T. L., Golshan, S., Jeste, D. V., & Gatz, M. (2004). Quality of life in geriatric generalized anxiety disorder:a preliminary investigation. *Journal of Psychiatric Research, 38*(3), 305–312. https://doi.org/10.1016/j.jpsychires.2003.09.003.
Loerinc, A. G. (2018). *Examining the effectiveness and feasibility of a self-guided version of positive affect treatment.* UCLA. Retrieved from https://escholarship.org/uc/item/9hj133gp.
Mennin, D. S., Ellard, K. K., Fresco, D. M., & Gross, J. J. (2013). United we stand: Emphasizing commonalities across cognitive-behavioral therapies. *Behavior Therapy, 44*(2), 234–248. https://doi.org/10.1016/j.beth.2013.02.004.
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.
Mitte, K. (2005). Meta-analysis of cognitive-behavioral treatments for generalized anxiety disorder: A comparison with pharmacotherapy. *Psychological Bulletin, 131*(5), 785.
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.* NICE. Retrieved from https://www.nice.org.uk/guidance/cg113.
Pew Research Center. (2019). *Demographics of mobile device ownership and adoption in the United States.* Pew Research Center: Internet, Science & Tech. Retrieved from https://www.pewinternet.org/fact-sheet/mobile/.
Qualtrics. (2019). *Qualtrics, version august 2019.* Provo, UT, USA. Retrieved from http://www.qualtrics.com.
R Core Team. (2018). R: A language and environment for statistical computing. Retrieved from https://cran.r-project.org/.
Reilly, M. C., Zbrozek, A. S., & Dukes, E. M. (1993). The validity and reproducibility of a work productivity and activity impairment instrument. *PharmacoEconomics, 4*(5), 353–365. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10146874.
Richards, D., & 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.
Richards, D., Richardson, T., Timulak, L., & McElvaney, J. (2015). The efficacy of internet-delivered treatment for generalized anxiety disorder: A systematic review and meta-analysis. *Internet Interventions, 2*(3), 272–282.
Richards, D., Timulak, L., Rashleigh, C., McLoughlin, O., Colla, A., Joyce, C., … Anderson-Gibbons, M. (2016). Effectiveness of an internet-delivered intervention for generalized anxiety disorder in routine care: A randomised controlled trial in a student population. *Internet interventions, 6*, 80–88.
Robinson, E., Titov, N., Andrews, G., McIntyre, K., Schwencke, G., & Solley, K. (2010). Internet treatment for generalized anxiety disorder: A randomized controlled trial comparing clinician vs. technician assistance. *PloS One, 5*(6), Article e10942. https://doi.org/10.1371/journal.pone.0010942.
Roemer, L., Orsillo, S. M., & Salters-Pedneault, K. (2008). Efficacy of an acceptance-based behavior therapy for generalized anxiety disorder: Evaluation in a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 76*, 1083–1089. https://doi.org/10.1037/a0012720.
Rounsaville, B. J., Carroll, K. M., & Onken, L. S. (2006). A stage model of behavioral therapies research: Getting started and moving on from stage I. *Clinical Psychology: Science and Practice, 8*(2), 133–142. https://doi.org/10.1093/clipsy.8.2.133.
Roy-Byrne, P. P., & Wagner, A. (2004). Primary care perspectives on generalized anxiety disorder. *Journal of Clinical Psychiatry, 65*(Suppl 13), 20–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15384933.
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. http://www.ncbi.nlm.nih.gov/pubmed/7967852.
Schermuly-Haupt, M. L., Linden, M., & Rush, A. J. (2018). Unwanted events and side effects in cognitive behavior therapy. *Cognitive Therapy and Research, 42*, 219–229. https://doi.org/10.1007/s10608-018-9904-y.
Shafran, R., Brosan, L., & Cooper, P. (2013). *The complete CBT guide for anxiety.* London: Constable & Robinson.
Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Nikles, J., Tate, R., et al. (2015). CONSORT extension for reporting N-of-1 trials (CENT) 2015: Explanation and elaboration. *BMJ, 350*(may14 15), h1793. https://doi.org/10.1136/bmj.h1793.
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. *Archives of Internal Medicine, 166*(10), 1092. https://doi.org/10.1001/archinte.166.10.1092.
Stapinski, L. A., Abbott, M. J., & Rapee, R. M. (2010). Evaluating the cognitive avoidance model of generalised anxiety disorder: Impact of worry on threat appraisal, perceived control and anxious arousal. *Behaviour Research and Therapy, 48*(10), 1032–1040. https://doi.org/10.1016/j.brat.2010.07.005.
Stein, M. B., & Heimberg, R. G. (2004). Well-being and life satisfaction in generalized anxiety disorder: Comparison to major depressive disorder in a community sample. *Journal of Affective Disorders, 79*(1–3), 161–166. https://doi.org/10.1016/S0165-0327(02)00457-3.
Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., et al. (2016). The single-case reporting guideline in BEhavioural interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology, 4*(1), 1–9. https://doi.org/10.1037/arc0000026.
Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., et al. (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.
Titov, N., Andrews, G., Johnston, L., Robinson, E., & Spence, J. (2010). Transdiagnostic internet treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy, 48*(9), 890–899. https://doi.org/10.1016/j.brat.2010.05.014.
Titov, N., Andrews, G., Robinson, E., Schwencke, G., Johnston, L., Solley, K., et al. (2009). Clinician-assisted internet-based treatment is effective for generalized

anxiety disorder: Randomized controlled trial. *Australian and New Zealand Journal of Psychiatry, 43*(10), 905–912. https://doi.org/10.1080/00048670903179269.

Wittchen, H.-U. (2002). Generalized anxiety disorder: Prevalence, burden, and cost to society. *Depression and Anxiety, 16*(4), 162–171. https://doi.org/10.1002/da.10065.

Wood, A. M., & Joseph, S. (2010). The absence of positive psychological (eudemonic) well-being as a risk factor for depression: A ten year cohort study. *Journal of Affective Disorders, 122*(3), 213–217. https://doi.org/10.1016/j.jad.2009.06.032.