## Title

Do the effects of internet-delivered cognitive-behavioral therapy (i-CBT) last after a year and beyond? A meta-analysis of 154 randomized controlled trials (RCTs).



## Authors

Zainal, Nur
Soh, Chui
Van Doren, Natalia
et al.

## Publication Date

2024-12-01

## DOI

10.1016/j.cpr.2024.102518

Peer reviewed

HHS Public Access

Author manuscript

*Clin Psychol Rev*. Author manuscript; available in PMC 2025 February 24.

Published in final edited form as:

*Clin Psychol Rev*. 2024 December ; 114: 102518. doi:10.1016/j.cpr.2024.102518.

# Do the effects of internet-delivered cognitive-behavioral therapy (i-CBT) last after a year and beyond? A meta-analysis of 154 randomized controlled trials (RCTs)

**Nur Hani Zainal**[a,b,*], **Chui Pin Soh**[b], **Natalia Van Doren**[c], **Corina Benjet**[d]

[a]Harvard Medical School, Department of Health Care Policy, Boston, MA, USA

[b]National University of Singapore, Department of Psychology, Kent Ridge Campus, Singapore

[c]University of California at San Francisco, Department of Psychiatry and Behavioral Sciences, San Francisco, CA, USA

[d]Center for Global Mental Health Research, National Institute of Psychiatry Ramón de la Fuente Muñiz, Mexico City, Mexico

## Abstract

Although the short-term efficacy of internet-delivered cognitive-behavioral therapy (i-CBT) is well-established, its long-term efficacy remains understudied. Robust variance estimation meta-analysis was thus conducted across guided and self-guided i-CBT, synthesizing data from 154 randomized controlled trials ($N$ = 45,335) with ≥ 12-month follow-ups. For binary outcomes, guided (52.3% vs. 38.6%; log-risk ratio [LOG-RR] = 1.15 95% confidence interval [1.04, 1.26]) yielded higher remission, reliable improvement, and response rates, and lower suboptimal treatment outcome rates (9.3% vs. 10.8%; LOG-RR = 0.63 [0.45, 0.80]) than treatment-as-usual, active controls, and waitlists at ≥12 months. Insufficient studies precluded testing the efficacy between self-guided i-CBT and controls for binary outcomes. For baseline-to-12-month dimensional outcomes, guided i-CBT produced greater reductions in anxiety, depressive, post-traumatic stress disorder (PTSD) symptoms, and repetitive negative thinking (Hedge's $g$ = −1.86 to −0.31), and self-guided i-CBT yielded stronger reductions in depressive symptoms ($g$ = −0.51) than all controls. For outcome scores aggregated at ≥ 12-month follow-ups, guided i-CBT alleviated anxiety, depression, distress, insomnia, PTSD symptoms, role impairment, emotion regulation, and quality of life ($g$ = −0.31 to 0.26), and self-guided i-CBT yielded lower anxiety and depressive symptoms ($g$ = −0.16 to −0.09) than all controls. No significant differences in efficacy emerged between guided and self-guided i-CBT when sufficient studies existed for a meta-analysis. There was no evidence for publication bias. Long-term efficacy was similar to short-term efficacy for most outcomes. Implementing scalable i-CBTs should entail transparency about their long-term benefits and drawbacks.

*Corresponding author at: Department of Health Care Policy, Harvard Medical School, Department of Psychology, National University of Singapore, Singapore. hanizainal@nus.edu.sg (N.H. Zainal).

Declaration of competing interest

None of the authors have conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2024.102518.

**Keywords**

Common mental disorders; Digital mental health interventions; Internet-delivered cognitive behavioral therapy; Long-term efficacy; meta-analysis; Randomized controlled trial

## 1. Introduction

Anxiety disorder, depression, obsessive-compulsive disorder (OCD), and post-traumatic stress disorder (PTSD) are among the most prevalent mental disorders (Fawcett et al., 2020; Georgieva et al., 2021) and carry significant personal and societal burdens (Baxter et al., 2014; Santomauro et al., 2021). Clinical guidelines advise both pharmacological and psychological treatments to remedy symptoms of these common mental disorders (CMDs; Bandelow et al., 2023; Leichsenring et al., 2023). However, the majority of clients prefer psychotherapy over psychotropic medications (McHugh et al., 2013). Of the various psychotherapies available, the one that has amassed the most evidence base for the strongest comparative efficacy for treatment remission, response, and alleviation of CMD symptoms is cognitive-behavioral therapy (CBT; Cuijpers et al., 2023).

CBT has been theorized to alleviate symptoms of anxiety, depression, OCD, and PTSD by reducing repetitive negative thinking (RNT) and unhelpful beliefs and reframing limiting self-talk and thoughts (Beck & Haigh, 2014). Beyond enhancing thought and emotion repertoires, CBT has been proposed to remedy CMD symptoms by decreasing emotionally driven, avoidant, compulsive, and related self-sabotaging behaviors (Tolin, 2016; Zainal et al., 2021; Zainal & Newman, 2019). Indeed, traditional face-to-face (FTF) CBT has accrued a large evidence base supporting its efficacy in treating CMDs (Cuijpers et al., 2023; Cusack et al., 2016; Reid et al., 2021; Springer et al., 2018). However, FTF CBT incurs notable financial, logistical, and time costs and is linked to attitudinal barriers such as shame and stigma (Clement et al., 2015; Goetter et al., 2020). Moreover, FTF CBT sessions entail being in the same physical location of a trained therapist and require more time from the clinician (Andrews et al., 2018).

Fortunately, testing and dissemination of scalable CBTs, especially internet-delivered CBTs (i-CBT), has been growing (Andrews & Williams, 2015). The term “Internet” in i-CBT encapsulates online computerized Web platforms and mobile phone or tablet apps, which users might opt to pay via subscriptions if accessed independently or receive free-of-charge if assigned to i-CBT by enrolling in a trial (Wasil et al., 2022). Note that i-CBT differs from general wellness apps, as i-CBT was explicitly developed for therapeutic purposes based on CBT principles (Andersson et al., 2019). Both i-CBT and traditional FTF CBT are grounded in the same theories, teaching behavioral activation, cognitive restructuring, exposure therapy, problem-solving, relaxation, relapse prevention, and related skills while delivering homework between modules or sessions (Andersson, 2016). These learned skills can be practiced in or generalized to diverse life contexts even after the online program ends to cope with stressors and symptoms effectively in self-reinforcing ways. Improved cognitive patterns could become more automatic or habitual across long durations, decreasing relapse probability by fruitfully enhancing and sustaining adaptive beliefs to manage life situations.

During the treatment phase (typically one to three months), users are regularly encouraged to apply these emotion regulation (ER) skills after the active i-CBT trial phase ends. After treatment ends, it is plausible that the changes in new lifestyles, mindsets, and routines promoted through i-CBT might promote lasting mental health and well-being improvements.

Regarding delivery format, i-CBTs are often supported by a supervised guide (non-specialized professional or layperson) or are fully self-guided (Simon et al., 2023). Fully self-guided i-CBTs are conducted without the assistance of a clinician, where these fully autonomous, low-intensity, self-guided formats involve no or only automated contact and rely mainly on intrinsic motivation for engagement and skill-building (Tong et al., 2024). Guided i-CBTs typically include remote human interaction via a communication platform. Human coaches or guides who interact with i-CBT users usually receive rigorous training and implementation fidelity supervision (Fitzsimmons-Craft et al., 2023).

Guidance within guided i-CBTs varies in intensity and structure (Andersson, 2024). Low-intensity, asynchronous, and unstructured guided on-demand support consists of a coach who only initiates contact in response to user inquiries about the i-CBT modules or an urgent clinical matter. Medium-intensity, asynchronous, and structured guidance features regular pre-planned 10 to 15-minute interactions via messaging platforms (Fitzsimmons-Craft et al., 2023). High-intensity, synchronous, and structured guidance comprises consistently scheduled audio phone check-ins or remote telehealth visits, ensuring reliable and real-time support (Krzyzaniak et al., 2024). Per ethical principles, asynchronous and synchronous interactions between the user and clinician or paraprofessional in guided i-CBT would occur on secure encrypted audio, email, messaging, or video platforms. These interactions aim to facilitate skill-building through accountability, encouragement, and feedback.

On a related note, low engagement is a key challenge in most digital mental health interventions during and after the intervention phase (Beatty & Binnion, 2016), and this problem is especially pronounced in self-guided i-CBTs (Baumeister et al., 2014). Numerous terms are considered interchangeable with i-CBT engagement (e.g., adherence, completion, uptake, usage) and tend to be defined as the number of logins, pages read, and modules completed (Beintner et al., 2019). Nonetheless, participants who were more actively involved in guided or self-guided i-CBT during and after treatment tended to reap greater mental health and functional outcome gains by continuing to practice the skills learned in i-CBT modules even one-year post-baseline (Eilert et al., 2023; Schure et al., 2020).

Despite challenges with engagement, numerous meta-analyses have shown that both guided and self-guided forms of i-CBTs outperformed controls in alleviating short-term clinical endpoints. These endpoints included symptoms of anxiety (Dulsen & Baumeister, 2024; Gutierrez et al., 2023; Johnson et al., 2024; Jonsson et al., 2023; Pauley et al., 2021; Wang et al., 2023), depression (Karyotaki et al., 2021; Kohnen et al., 2021; Linardon et al., 2024; Lo et al., 2024; Tong et al., 2023; Xiong et al., 2023), other CMD symptoms such as insomnia, OCD, and PTSD (Bai et al., 2024; Hasanvandi et al., 2022; Lewis et al., 2019; Polak & Tanzer, 2024; Tng et al., 2024; Zhang et al., 2023) as well as related outcomes such as

distress and self-efficacy (Fernandez-Rodriguez et al., 2024). Another meta-analysis showed that guided i-CBT fared better than self-guided i-CBT at yielding larger anxiety symptom severity reductions during posttreatment, although such differential efficacy disappeared at follow-ups (Oey et al., 2023). A common denominator of these meta-analyses was that the timeframe of baseline to follow-up data primarily ranged from three to six months. However, there is limited meta-analytic evidence regarding efficacy in the longer term. Long-term efficacy was defined herein as ≥ 12 months post-randomization, as most i-CBT studies capped their long-term follow-up at the time point.

Moreover, accruing long-term evidence for the efficacy of i-CBT after at least one-year post-randomization is essential, as CMD symptoms commonly follow a chronic course that stretches several years (Kuhne et al., 2020; Steinert et al., 2015; Struijs et al., 2018). One school of thought posits that short-term intervention would not be sufficient to preserve i-CBT enhancements (Clark, 2018; Vittengl et al., 2007). Comparatively, another theoretical perspective asserts that i-CBT could confer both immediate mental health symptom alleviation and sustained benefits after 12 months and beyond (Nakao et al., 2021). Plausibly, participants would still be applying skills learned from i-CBT modules to change engrained patterns of behaviors, thinking, and lifestyles that no longer serve their goals and values (Dalle Grave et al., 2020; Saether et al., 2019). The plethora of coping and ER strategies learned in i-CBT might continue to effectively manage symptoms and promote quality of life (QOL) even after the intervention has concluded (Svardman et al., 2022). Together, further mental health improvements in the post-intervention phase of the trial are theoretically possible.

However, little is known about whether guided vs. self-guided i-CBT achieves similar efficacy over the short- and long-term. Determining any differences in delivery mechanism is essential as guided i-CBT is more costly and burdensome, whereas self-guided i-CBT may have broader reach and accessibility. Enduring, long-term efficacy evidence is crucial for informing treatment research priorities, such as further exploration of treatment response factors and optimization strategies, and for clinicians to provide clients with pragmatic expectations.

Four recent meta-analyses have explored the long-term effectiveness of i-CBT on anxiety, depression, and related CMD symptoms. Andersson et al. (2018) aggregated data across 14 studies and found a large decrease in symptoms persisting for up to five years from baseline. However, without a control, this meta-analysis assessed i-CBT effectiveness solely based on its progression over time (from baseline to posttreatment to follow-up). Hence, it could not differentiate i-CBT effectiveness from naturalistic remission or potential placebo responses. Overcoming this limitation, Sztein et al. (2018) and Mamukashvili-Delau et al. (2023) synthesized 14 to 15 randomized controlled trials (RCTs) of i-CBTs and observed medium-to-large effect sizes of i-CBT over controls. However, these meta-analyses solely examined depression outcomes and included only one RCT when analyzing long-term outcomes defined as > 8 months post-baseline. In addition, Lo et al. (2023) found that i-CBT had moderate effects on anxiety and depression symptom severity after six months or more but grouped RCTs with non-RCTs, precluding stringent causal inferences. In addition, most of these i-CBT trials included in existing meta-analyses with ≥ 12-month follow-ups were

conducted in Europe and North America, with countries such as Sweden showing exemplary foresight and leadership through early adoption and integration of digital health treatments into their healthcare systems (Brantnell et al., 2020). An up-to-date quantitative synthesis is thus needed to inform optimal approaches to maximize the long-term efficacy of i-CBT, such as cross-cultural adaptations, and build on existing qualitative reviews (Naderbagi et al., 2024).

Thus, we meta-analyzed data from i-CBT studies across the globe to test the long-term efficacy of i-CBT. Both guided and self-guided i-CBTs were compared to active, FTF, treatment-as-usual (TAU), or waitlist (WL) control to test their efficacy on anxiety, depression, OCD, and PTSD symptoms, as well as distress, QOL, role impairment, and related outcomes at ≥12 months post-randomization (our definition of long-term effects). We aimed to fill knowledge gaps in several ways. First, we comprehensively tested if i-CBT was superior to active, TAU, and WL controls on long-term clinically relevant outcomes. The first aim was to compare guided and self-guided i-CBT with other controls and to contrast both forms of i-CBT with FTF CBT. Second, we evaluated how long-term efficacy outcomes compared to short-term efficacy outcomes. Third, it remains unclear for whom i-CBT would be more efficacious over controls in the long term. Therefore, we conducted exploratory meta-regression and subgroup analyses to test how various clinical, socio-demographic, treatment, and study-specific variables modified the efficacy of i-CBT on stated outcomes.

## 2. Method

### 2.1. Eligibility criteria and procedures

The meta-analysis followed the criteria outlined in the PRISMA statement (Moher et al., 2015) and was preregistered (Open Science Framework [OSF] [https://osf.io/k4a53] and PROSPERO [CRD42024519989]). We methodically sought eligible RCTs across various databases: Cochrane Library, EMBASE, PsycINFO, PubMed, ScienceDirect, and Web of Science. Grey literature (i.e., dissertations, preprints, and theses via Google Scholar and OSF) was also consulted. Combinations and subsets of these search terms were used: “12 months,” “anhedonia,” “anxiety,” “app,” “CBT,” “cognitive-behavioral therapy,” “computerized,” “depression,” “digital,” “generalized anxiety,” “Internet,” “long term,” “major depressive disorder,” “online,” “OCD,” “obsessive-compulsive disorder,” “panic,” “perseverative negative thinking,” “persistent depressive disorder,” “post-traumatic stress,” “PTSD,” “randomized controlled trial,” “RCT,” “rumination,” “smart-phone,” “treatment-as-usual,” “waiting list,” “Web,” and “worry.” The inclusion criteria included an RCT design with active, TAU, or WL controls, evaluation of guided or self-guided i-CBT, long-term follow-up at 12 months or more, and symptom severity outcomes of anxiety, depression, OCD, or PTSD. The exclusion criteria comprised non-RCT designs, evaluation of non-internet-delivered CBTs (e.g., bibliotherapy, FTF, telephone), other internet-delivered psychotherapies (e.g., family systems therapy, psychodynamic, supportive listening), absence of a follow-up at 12 months or more, and no measure of anxiety, depression, OCD, or PTSD symptom outcomes. Although the English language was not an inclusion criterion, we did not find any eligible articles written in other languages. For the search duration,

potentially eligible articles published between January 2004 and August 2024 were screened before the inclusion of eligible studies.

### 2.2. Data extraction

Two researchers (NHZ, SCP) independently gathered data from all the studies included and recorded it in a database. Two other authors (NVD, CB) verified the accuracy of this database through cross-checking. The summary of variables comprised primary author and publication year, nation and continent of RCT, age (mean ± *SD* and range), gender distribution (% women), condition-specific sample size, duration of i-CBT, and modality (i.e., guided vs. self-guided). Also, it encompassed the content of the control group (WL, TAU, or active controls), wherein active controls included online forums, physical exercise, psychoeducation, self-monitoring, and others. FTF CBT was defined as in-person CBT delivered in a group or individually. We also extracted data about the outcome, the measures (total scores and subscales), and attrition. Further, we collected data on concurrent medication (percentage of pharmacotherapy use) and prior psychotherapy. The means and *SD*s of symptom severity outcome measures were collected during the baseline, posttreatment, and follow-up assessments. Positive treatment outcomes (remission, response, reliable improvement, and reliable recovery) and suboptimal treatment outcomes (antidepressant medication [ADM] or other pharmacotherapy use, relapse, sick leave, treatment utilization, and related variables) in binary formats were also extracted. If the control was WL, we extracted the earlier data points from the original pre-post analysis and used multiple imputation to impute follow-up scores collected at 12 months or beyond post-randomization if such data were missing or extracted the reported follow-up scores if such data were available. Appendix 1 summarizes the 154 RCTs in the present meta-analysis, including more information about the TAU control. Appendix 2 summarizes the measures classified under specific outcomes. Appendix 3 offers a concise description of i-CBT core modules.

### 2.3. Meta-analysis of comparative efficacy of i-CBT on long-term outcomes

All analyses were conducted using *R* software (R Core Team, 2024), and the analytic scripts are summarized in Appendix 4. To handle missing data, multiple imputation was conducted using the *missRanger* package (Mayer, 2024). Studies showed that even when high rates of missing data were imputed, the resulting grand means and confidence intervals (CIs) closely resembled those obtained via fully informed weighted analyses (Kambach et al., 2020), generating valid estimates with nested data (Wijesuriya et al., 2020).

Given likely variations in effect sizes across studies, random-effects modeling with restricted maximum likelihood estimation was employed for all outcomes with the *metafor* package (Viechtbauer, 2010). Sensitivity analyses were conducted using two-level random effects (Cheung, 2008), three-level random effects (Cheung, 2014), and robust variance estimation (RVE; Tipton & Pustejovsky, 2015) meta-analysis, also with the *metafor* package. Whereas two-level random-effects meta-analysis adjusts for within-study heterogeneity, three-level random-effects meta-analysis accounts for both within- and between-study heterogeneity. However, neither two-level nor three-level random-effects meta-analyses account for correlated sampling errors arising from multiple effect sizes nested within each study. RVE

resolves these issues by handling such effect size dependencies using sandwich estimators (Tipton, 2013; Tipton & Pustejovsky, 2015). Results from RVE meta-analyses thus were reported, and other sensitivity analyses are available in Appendix 5. All models were compared across guided vs. self-guided i-CBT.

For binary outcomes, we calculated log risk ratios (LOG-RR) coupled with 95% CIs. For continuous outcomes, we computed controlled effect sizes to assess the difference between i-CBT and control conditions in outcomes (comparative efficacy), utilizing sample size bias-corrected Hedge's *g* and the corresponding 95% CIs (Hedges et al., 2010). The mean pre-follow-up change scores were subtracted between groups and divided by the pooled *SD*s from both groups. *SD*s were computed using recommended formulas (Morris & DeShon, 2002). Since pre-follow-up correlations of continuous scores were not obtainable, a conservative estimate of 0.50 was used. Hedge's *g* values of 0.2, 0.5, and 0.8 signified small, moderate, and large effect sizes (Cohen, 1988). Alpha correction was applied (Simes, 1986). To maximize power and precision and minimize biases and errors in parameter estimates, all meta-analyses and interpretations were conducted under the condition of five or more studies ($k \geq 5$; Deeks et al., 2023).

### 2.4. Meta-regression to explore heterogeneous treatment effects (HTEs)

Expanding upon prior research, we anticipated notable HTE (Terhorst et al., 2024) and thus employed a random-effects model to accommodate the heterogeneity among the included studies. We assessed effect size heterogeneity using the $I^2$ and $Q$ statistics. The $I^2$ statistic varies from 0 to 100%, where 25% signifies low, 50% moderate, and 75% high heterogeneity (Higgins & Thompson, 2002). A $Q$ statistic $p$-value below .05 suggests the presence of heterogeneity (Cochran, 1954). Outliers were identified as studies with a 95% CI that did not overlap with the overall effect size's CI (Viechtbauer & Cheung, 2010). Outliers were also first Winsorized by replacing them with the 2.5% or 97.5% values (Hoo et al., 2002), and the remaining outliers were then deleted (Viechtbauer & Cheung, 2010) using the *dplyr* (Wickham et al., 2023) and *psych* (Revelle, 2024) packages. Influence tests were performed via standard diagnostic statistics (Cook's distance, difference in fits [DFFITS], hat values, and study weights). For categorical moderators (i.e., effect modifiers), subgroup analyses were conducted. For continuous moderators, our meta-regression analyses employed a restricted maximum likelihood model, integrating the Knapp–Hartung method (Borenstein et al., 2009) to account for potential biases and uncertainties in the estimation of effect sizes and their standard errors.

### 2.5. Risk of bias evaluation

We evaluated the study quality utilizing the Cochrane Risk of Bias tool 2 (RoB 2.0; Sterne et al., 2019) and the *robviz* package (McGuinness, 2019). Selection bias (including allocation concealment and allocation sequence generation), performance bias (involving deviations from intended intervention protocols), detection bias (comprising blinding of outcome evaluators), attrition bias (related to missingness), and reporting bias (involving selective outcome reporting) were evaluated. Two authors (NHZ, SCP) independently categorized the trials based on specific risk of bias criteria. Differences were settled with the involvement of a third and senior author (NVD, CB). RCTs deemed to have low risk of bias across

all domains were categorized as trials with an overall low risk of bias. In addition, we generated contour-enhanced funnel plots to test publication bias or other potential small-study effects when comparing i-CBT vs. controls (Duval & Tweedie, 2000) and conducted the rank-correlation test (Begg & Mazumdar, 1994) using the *metafor* package.

## 3. Results

### 3.1. Inter-rater agreement

Significantly satisfactory-to-excellent levels of inter-rater agreement (Koo & Li, 2016) were found for categorical (weighted Cohen's $\kappa$ = .91, range = .81 to 1.00) and dimensional variables (mean intra-class correlation coefficient [ICC] = .94, range = .65 to 1.00).

### 3.2. Baseline variables

Fig. 1 shows the PRISMA flowchart. A total of 154 RCTs were included in our meta-analysis. The total sample size was 45,335, averaging 294 per study, which ranged from 19 to 6386. Age averaged 40.82 years (*SD* = 13.98, range = 9.97 to 74.46). Women averaged 65.7%, ranging from 0% to 100%. Retention percentage averaged 71.2%, ranging from 10% to 100%. The number of core i-CBT modules averaged 8.42 (*SD* = 3.38, range = 3 to 30). The number of weeks of the intervention phase averaged 10.06 (*SD* = 5.00, range = 3 to 36). At baseline, an average of 54.3% had prior psychotherapy, whereas 38.9% were taking psychotropic medications, including ADM. Most RCTs took place in *Europe* (Denmark: 3/154 [2.0%]; Finland: 2/154 [1.3%]; Germany: 18/154 [11.7%]; Spain: 6/154 [3.9%]; Sweden: 48/154 [31.2%]; Switzerland 1/154 [0.6%]; The Netherlands: 25/154 [16.2 %]; The United Kingdom (UK): 12/154 [7.8 %]), followed by *Australia* (21/154 [13.6%]), *North America* (Canada: 3/154 [2.0%]; The United States of America (USA): 9/154 [5.8%]), and *Asia* (China: 2/154 [1.3%]; Japan: 3/154 [2.0%]). Regarding crude proxies of socioeconomic status (SES), the mean proportion of participants with higher education was 51.4%, and those who were gainfully employed comprised 48.5% of the total sample. Five studies reported on income. Of these, 62.8% of participants had an annual income of ≥ USD 30,000.

### 3.3. Predicting CMD remission, reliable improvement, and response outcomes at ≥ 12 months

**Comparison with all controls.—**First, we evaluated the efficacy of remission rates. Guided i-CBT yielded significantly higher remission rates than all controls (1591/2683 [59.3%] vs. 974/2496 [39.0%], number of studies [*k*] = 18, log-risk ratio [LOG-RR] = 1.20, 95% confidence interval [CI] [1.04, 1.36], $p < .001$, $\tau^2 = 0.07$ [standard error] [0.04]). However, insufficient studies ($k = 3$) hindered testing the efficacy between self-guided i-CBT and all controls on remission rates.

Second, we tested the efficacy of reliable improvement rates. Guided i-CBT yielded significantly higher reliable improvement rates than all controls (1043/2258 [46.2%] vs. 875/2316 [37.8%], $k = 20$, LOG-RR = 1.07 [0.99, 1.16], $p < .001$, $\tau^2 = 0.00$ [0.01]). Nonetheless, inadequate studies ($k = 2$) precluded testing the efficacy between self-guided i-CBT and all controls on reliable improvement rates.

Third, we assessed the efficacy of treatment response rates. Guided i-CBT yielded significantly higher response rates than all controls (512/1222 [41.9%] vs. 462/1221 [37.8%], $k = 10$, LOG-RR = 1.02 [0.81, 1.23], $p < .001$, $\tau^2 = 0.05$ [0.05]). No RCTs have tested response rate differences between self-guided i-CBT and any control.

**Comparison with distinct controls.—**Compared to TAU, rates of positive treatment outcomes examined above were significantly higher in guided i-CBT (1366/2455 [55.6%] vs. 893/2356 [37.9%], $k = 9$, LOG-RR = 1.32 [1.07, 1.57], $p < .001$, $\tau^2 = 0.09$ [0.07]; Fig. 2). Similarly, positive treatment outcome rates were significantly higher in guided i-CBT than in active controls (799/1645 [48.6%] vs. 700/1562 [44.8%], $k = 17$, LOG-RR = 1.03 [0.92, 1.15], $p < .001$, $\tau^2 = 0.03$ [0.02]). However, there were insufficient RCTs to compare guided with self-guided i-CBT ($k = 3$) as well as self-guided i-CBT with TAU ($k = 3$) and active controls ($k = 1$).

### 3.4. Predicting suboptimal treatment outcomes at ≥ 12 months

Suboptimal treatment outcomes comprised reliable deterioration, incident relapse in anxiety, depression, or OCD, use of psychotropic medications, harmful alcohol or substance use, treatment utilization, or sick leave. These outcomes were aggregated instead of examined separately, given the few RCTs. Suboptimal treatment outcome rates were significantly lower in guided i-CBT compared to all controls (384/4149 [9.3%] vs. 416/3863 [10.8%], $k = 17$, LOG-RR = 0.63 [0.45, 0.80], $p < .001$, $\tau^2 = 0.05$ [0.05]; Fig. 3). Specifically, compared to TAU, suboptimal treatment outcome rates were significantly lower in guided i-CBT (158/1246 [12.7%] vs. 219/1232 [17.8%], $k = 7$, LOG-RR = 0.65 [0.42, 0.87], $p < .001$, $\tau^2 = 0.06$ [0.05]). Similarly, compared to active controls, suboptimal treatment outcome rates were significantly lower in guided i-CBT (173/1372 [12.6%] vs. 177/1283 [13.8%], $k = 9$, LOG-RR = 0.81 [0.47, 1.14], $p < .001$, $\tau^2 = 0.15$ [0.13]). There were insufficient RCTs to meta-analyze the effect of self-guided i-CBT vs. all controls ($k = 2$), self-guided i-CBT vs. active controls ($k = 2$), and guided i-CBT vs. self-guided i-CBT ($k = 1$) on suboptimal treatment outcomes at ≥ 12 months.

### 3.5. Long-term efficacy of guided i-CBT on dimensional CMD symptoms and related outcomes

**Predicting baseline-to-12-month changes.—**Consistent with predictions, guided i-CBT was significantly more efficacious than all controls in reducing baseline-to-12-month depression ($k = 81$, $g = -0.31$ [−0.44, −0.18], $p < .001$), anxiety ($k = 44$, $g = -0.44$ [−0.58, −0.31], $p < .001$), RNT ($k = 10$, $g = -0.69$ [−1.05, −0.33], $p < .001$), and PTSD symptom severity ($k = 5$, $g = -1.86$ [−2.37, −1.36], $p < .001$; Table 1a and Appendix 5 Table S1). Contrary to predictions, no significant differences in efficacy emerged between guided i-CBT and all controls on baseline-to-12-month change in role impairment, ER, insomnia, global symptoms, pain, QOL, and self-efficacy. There were insufficient RCTs to test treatment efficacy on anxiety sensitivity as well as symptoms of OCD, alcohol, or substance use. Compared to active controls, guided i-CBT was significantly more efficacious in decreasing baseline-to-12-month depression ($k = 31$, $g = -0.34$ [−0.60, −0.08], $p = .011$) and anxiety ($k = 21$, $g = -0.36$ [−0.70, −0.02], $p = .039$; Appendix 5 Table S2). Compared to TAU, guided i-CBT was significantly more efficacious in reducing baseline-to-12-month

depression ($k$ = 27, $g$ = −0.40 [−0.59, −0.21], $p$ < .001), anxiety ($k$ = 6, $g$ = −0.61 [−0.79, −0.43], $p$ < .001), and role impairment ($k$ = 12, $g$ = −0.24 [−0.44, −0.04], $p$ = .018; Appendix 5 Table S3). No significant differences in efficacy emerged between guided and self-guided i-CBT on baseline-to-12-month change in any outcomes (Appendix 5 Table S4).

**Predicting scores aggregated at ≥ 12 months.—**The outcome of this set of analyses was derived by pooling the means and *SDs* of the scores reported at follow-ups of ≥ 12 months after randomization instead of examining pre-follow-up change in scores as with the prior analyses. Consistent with predictions, guided i-CBT yielded significantly lower depression ($k$ = 113, $g$ = −0.18 [−0.25, −0.12], $p$ < .001), anxiety ($k$ = 78, $g$ = −0.23 [−0.30, −0.15], $p$ < .001), distress ($k$ = 31, $g$ = −0.12 [−0.21, −0.03], $p$ = .006), role impairment ($k$ = 41, $g$ = −0.18 [−0.29, −0.08], $p$ = .001), insomnia ($k$ = 16, $g$ = −0.31 [−0.46, −0.16], $p$ < .001), and PTSD ($k$ = 8, $g$ = −0.28 [−0.42, −0.14], $p$ < .001) scores as well as higher QOL ($k$ = 74, $g$ = 0.11 [0.06, 0.17], $p$ < .001) and ER ($k$ = 12, $g$ = 0.26 [0.14, 0.38], $p$ < .001) scores than all controls at ≥ 12 months (Table 1b and Appendix 5 Table S5). Contrary to predictions, no significant differences in efficacy emerged between guided i-CBT and all controls on scores at ≥ 12 months for the remaining outcomes: distress, RNT, global symptoms, self-efficacy, pain, anxiety sensitivity, OCD, and alcohol or substance use. Compared to active controls, guided i-CBT did not yield differential efficacy significantly on any of these outcomes at ≥ 12 months (Appendix 5 Table S6). Compared to TAU, guided i-CBT yielded significantly lower depression ($k$ = 17, $g$ = −0.33 [−0.40, −0.27], $p$ < .001), anxiety ($k$ = 12, $g$ = −0.26 [−0.46, −0.06], $p$ = .011), role impairment ($k$ = 15, $g$ = −0.18 [−0.30, −0.07], $p$ = .002), and QOL scores ($k$ = 9, $g$ = −0.21 [−0.30, −0.12], $p$ < .001) as well as significantly higher distress scores at ≥ 12 months ($k$ = 22, $g$ = 0.11 [0.02, 0.19], $p$ = .016; Appendix 5 Table S7). On all outcomes, no significant comparative efficacy emerged between guided i-CBT and self-guided i-CBT (Appendix 5 Table S8).

### 3.6. Long-term efficacy of self-guided i-CBT on dimensional CMD symptom and related outcomes

**Predicting baseline-to-12-month changes.—**As hypothesized, self-guided i-CBT was significantly more efficacious than all controls in reducing baseline-to-12-month depression ($k$ = 23, $g$ = −0.51 [−0.88, −0.14], $p$ = .007), but not anxiety and role impairment (Table 2a and Appendix 5 Table S9). Compared to active controls, self-guided i-CBT was significantly more efficacious in reducing baseline-to-12-month depression ($k$ = 13, $g$ = −0.29 [−0.53, −0.05], $p$ = .018) but not anxiety (Appendix 5 Table S10). Contrary to hypotheses, no significant differences in efficacy emerged between self-guided i-CBT and TAU (Appendix 5 Table S11).

**Predicting scores aggregated at ≥ 12 months.—**As hypothesized, self-guided i-CBT yielded significantly lower scores of anxiety ($k$ = 12, $g$ = −0.16 [−0.23, −0.09], $p$ < .001) and depression ($k$ = 14, $g$ = −0.09 [−0.14, −0.04], $p$ = .001) than all controls at ≥ 12 months (Appendix 5 Table S12). Contrary to hypotheses, no significant differences in efficacy emerged between self-guided i-CBT and all controls on scores at ≥ 12 months for anxiety severity and role impairment. Compared to active controls, self-guided i-CBT yielded significantly lower depression scores ($k$ = 14, $g$ = −0.16 [−0.29, −0.02], $p$ = .021; Appendix

5 Table S13). Compared to TAU, self-guided i-CBT was significantly more efficacious in decreasing follow-up depression scores ($k = 5$, $g = -0.06$ [−0.13, −0.01], $p = .042$; Appendix 5 Table S14). Tables S15 to S18 in Appendix 5 present efficacy outcomes compared to WL. Sensitivity analyses suggested a similar pattern of findings for all examined treatment efficacy effects on binary and dimensional outcomes.

### 3.7. Exploring the efficacy of guided/self-guided i-CBT vs. FTF CBT

Twenty-three RCTs compared guided/self-guided i-CBT with FTF CBT. Guided/self-guided i-CBT was significantly more efficacious than FTF CBT in reducing baseline-to-12-month anxiety severity ($k = 12$, $g = -0.49$ [−0.87, −0.12], $p = .010$), but not depression, QOL, and role impairment (Appendix 5 Table S19). Similarly, guided/self-guided i-CBT produced significantly lower anxiety severity scores aggregated at all long-term follow-ups than FTF CBT ($k = 14$, $g = -0.25$ [−0.46, −0.04], $p = .019$), but no differential efficacy was observed for depression, QOL, and role impairment (Appendix 5 Table S20).

### 3.8. Meta-regression analyses to explore HTEs

We examined baseline-to-12-month changes in these outcomes in our HTE analyses: depression, anxiety, QOL, role impairment, insomnia, RNT, global symptoms, self-efficacy, ER, PTSD, pain, and alcohol or substance use. Due to space constraints, only significant meta-regression analyses are reported in the main manuscript (Appendix 5 Tables S21 to S36).

**Follow-up duration.—**The efficacy of guided i-CBT vs. all controls was significantly weaker in the long term than in the short term for insomnia ($k = 15$, $\beta = -0.29$ [−0.52, −0.06], $p = .014$). In contrast, longer follow-up duration was associated with significantly stronger efficacy of guided i-CBT vs. all controls on ER ($k = 12$, $\beta = 0.76$ [0.35, 1.17], $p < .001$), and QOL ($k = 78$, $\beta = 0.27$ [0.11, 0.43], $p = .001$; Appendix 5 Table S21). However, follow-up duration did not significantly moderate the efficacy of self-guided i-CBT vs. all controls (Appendix 5 Table S22).

**Age.—**Older age was related to significantly weaker efficacy of guided i-CBT vs. all controls on reduction in anxiety ($k = 76$, $\beta = -0.02$ [−0.03, −0.01], $p = .028$) and depression ($k = 108$, $\beta = -0.01$ [−0.03, −0.01], $p = .017$; Appendix 5 Table S23). Comparatively, age did not significantly moderate the effect of self-guided i-CBT vs. all controls (Appendix 5 Table S24).

**Gender.—**Gender (defined as the percentage of women) did not significantly moderate the comparative effect of guided i-CBT vs. all controls (Appendix 5 Table S25). However, higher percentage of women was associated with significantly stronger efficacy of self-guided i-CBT vs. all controls on QOL ($k = 5$, $\beta = 0.01$ [0.01, 0.03], $p = .039$; Appendix 5 Table S26).

**Attrition.—**Attrition (indexed as retention percentage at the last follow-up) did not significantly moderate the efficacy of guided i-CBT vs. all controls (Appendix 5 Table S27) and self-guided i-CBT vs. all controls (Appendix 5 Table S28).

**Core modules.—**Details on how each study defined core i-CBT modules (i.e., specific CBT skills included in each module) are provided in Appendix 3. Number of core modules did not significantly modify the efficacy of guided i-CBT vs. all controls (Appendix 5 Table S29). However, higher number of core modules was significantly associated with stronger efficacy of self-guided i-CBT vs. all controls on QOL ($k = 5$, $\beta = 0.09$ [0.02, 0.16], $p = .013$; Appendix 5 Table S30).

**Weeks of treatment.—**More weeks of treatment were significantly associated with larger efficacy of guided i-CBT vs. all controls on anxiety severity ($k = 76$, $\beta = 0.05$ [0.01, 0.09], $p = .015$; Appendix 5 Table S31). In addition, more weeks of treatment were significantly associated with greater efficacy of self-guided i-CBT vs. all controls on QOL ($k = 5$, $\beta = 0.02$ [0.01, 0.03], $p = .030$; Appendix 5 Table S32).

**SES.—**Higher education was not a significant moderator of the efficacy of guided i-CBT (Appendix 5 Table S33) or self-guided i-CBT (Appendix 5 Table S34). Full-time employment also did not significantly moderate the efficacy of guided i-CBT (Appendix 5 Table S35) or self-guided i-CBT (Appendix 5 Table S36).

### 3.9. Predicting retention at the last follow-up of ≥ 12 months

Appendix 5 Table S37 summarizes how various i-CBT comparisons predicted retention percentage as a proxy marker of intervention engagement. Guided i-CBT produced significantly lower retention rates than all controls ($k = 146$, LOG-RR = −0.05 [−0.08, −0.02], $p < .001$), active controls ($k = 48$, LOG-RR = −0.05 [−0.09, −0.01], $p = .023$), and TAU ($k = 49$, LOG-RR = −0.08 [−0.14, −0.03], $p = .003$), but not FTF CBT. Self-guided i-CBT also led to significantly lower retention percentage than all controls ($k = 33$, LOG-RR = −0.06 [−0.12, −0.01], $p = .050$). Conversely, guided i-CBT yielded significantly higher retention rates than self-guided i-CBT ($k = 17$, LOG-RR = 0.14 [0.03, 0.25], $p = .029$). There were insufficient studies to compare self-guided i-CBT and FTF CBT.

### 3.10. Study quality (Risk of bias evaluation)

For selection bias, most studies showed low risk of bias (107/154 [69.5%]), followed by some concerns (31/154 [20.1%]), and high risk of bias (16/154 [10.4%]; Fig. 4a). For performance bias, a similar risk of bias pattern was observed (low: 121/154 [78.6%]; some concerns: 20/154 [13.0%]; high: 13/154 [8.4%]). For attrition bias, most studies had low risk of bias (125/154 [81.2%]), whereas the remaining had high risk of bias (29/154 [18.8%]). For detection bias, most studies had low risk of bias (120/154 [77.9%]), whereas the rest had high risk of bias (34/154 [22.1%]). For reporting bias, most studies had low risk of bias (122/154 [79.2%]), whereas the remaining showed high risk of bias (32/154 [20.8%]). Collectively, about half of the studies had low risk of bias (81/154 [52.6%]), followed by high risk of bias (51/154 [33.1%]) and some concerns (22/154 [14.3%]). In addition, all rank-correlation tests to identify publication bias were insignificant for all dimensional (Appendix 5 Table S38) and binary outcomes (Appendix 5 Table S39). Moreover, all funnel plots appeared symmetrical (Fig. 4b and c depict overall plots).

## 4. Discussion

The present study meta-analyzed the long-term efficacy of guided and self-guided i-CBTs on CMD outcomes. Data from 154 RCTs ($N$ = 45,335) across 15 nations were evaluated, and long-term outcomes from 1 to 6 years post-randomization were examined. Plausible clinical, theoretical, and policy implications are discussed to advance clinical science.

Notably, remission, reliable improvement, and response rates for anxiety, depression, OCD, and PTSD were significantly higher for guided relative to all controls (52.3% vs. 38.6%), 12 to 72 months post-randomization. Suboptimal treatment outcome rates were also lower for guided i-CBT than all controls (9.3% vs. 10.8%). Similar patterns were observed when active, TAU, and WL controls were tested separately. These results suggested that guided i-CBT shows long-term efficacy in terms of remission, reliable improvement, and treatment response. Further, more studies are needed to evaluate self-guided modalities in terms of these categorical outcomes since there were insufficient RCTs to compare positive and suboptimal outcome rate differences between self-guided i-CBTs and controls.

Other noteworthy findings emerged when testing baseline-to-12-month change outcomes. Compared to all controls, guided i-CBT generated more reduction in anxiety, depression, PTSD severity, and RNT, and self-guided i-CBT yielded larger decreases in depression severity from baseline to 12 months. By instructing users to complete modules teaching various coping, ER, and graduated exposure skills sequentially and flexibly in diverse settings, both i-CBT types could sustain mental health enhancements (Bai, 2023; McCall et al., 2023).

Relatedly, the efficacy of guided/self-guided i-CBT did not differ from FTF CBT on all outcomes, except that guided/self-guided i-CBT was more effective than FTF CBT in decreasing anxiety severity. This observation could be partly accounted for by the more versatile delivery of i-CBTs than FTF CBT, which could decrease anxiety severity by reinforcing behavioral activation, exposure, and relaxation exercises in real-time (Kumar et al., 2017). For other outcomes, our findings paralleled other meta-analysis data that guided/self-guided i-CBT was equally clinically effective as and even more cost-effective than FTF CBT (Luo et al., 2020), rendering it a viable option for clients and clinicians who prefer it. Further, our results reinforce meta-analytic evidence that self-guided i-CBT is far from harmful (Karyotaki et al., 2018) and may serve as an initial therapy for anxiety and depression, offering an alternative to a vigilant waiting strategy in time-limited primary care contexts and other resource-constrained settings delivering FTF CBT. Future research should also extend these questions into OCD and PTSD outcomes, given the dearth of i-CBT RCTs.

Simultaneously, examining dimensional scores aggregated across follow-up assessment periods at ≥12 months after randomization revealed other favorable outcomes. Relative to all controls, both guided and self-guided i-CBT were consistently associated with fewer long-term anxiety and depression symptoms. Further, guided i-CBT yielded lower distress, role impairment, insomnia, and PTSD symptom scores, as well as higher ER and QOL in the long term. Encouragingly, these findings concurred with the notion that guided i-CBT shows promise as a tool for enhancing under-investigated functional outcomes for clients with

CMDs (Maj et al., 2023). Functional outcome constructs, such as ER, are multidimensional (Daros et al., 2021), and guided i-CBTs might have helped clients with CMD symptoms to use ER skills to optimally engage (or disengage) across various situations (Moltrecht et al., 2021). The positive role of a guide to potentially facilitate ER skills during exposure to and processing of trauma-related memories and situations head-on and encourage meaning-making in i-CBTs remains subject to debate (Kuester et al., 2016). Future empirical studies should probe these inquiries.

Inconsistent meta-analytic outcomes pertained to how guided i-CBT generated higher distress and lower QOL scores at follow-up than TAU. Such discrepancies might arise because meta-analysis examining pooled follow-up scores did not control for pre-treatment scores. These inconsistencies also imply that the comparative effects of guided i-CBT on distress and QOL endpoints reflect literature gaps warranting further attention, as similarly noted by other systematic reviews (e.g., Adhikary et al., 2023).

Our meta-regression analyses could account for only a portion of HTEs in the context of i-CBTs. In contrast to a meta-analysis that found no moderating effect of follow-up duration on the efficacy of digital mental health interventions (Moshe et al., 2021), our meta-regressions implied that the efficacy of guided i-CBT decayed over time for insomnia yet strengthened for ER and QOL. This discrepancy was probably due to the briefer follow-up durations in Moshe et al. (2021)’s meta-analysis. The absence of the moderating effect of follow-up duration on other examined outcomes (anxiety, depression, pain, PTSD symptoms, ER, and RNT) across time signals the necessity for strategies to sustain long-term i-CBT engagement (van Ballegooijen et al., 2014). Our results add essential data on the longer-term efficacy of i-CBTs, implying that booster sessions promoting prolonged therapy skills usage could maintain short-term benefits by consolidating CBT skills acquired during the trial intervention stages (Hadjistavropoulos et al., 2022; Peynenburg et al., 2022). More research is needed to devise feasible methods to sustain the initial benefits of i-CBT in the long run to assist people with CMD symptoms in alleviating the intensity of their symptoms or in bridging the delayed-access gap until they obtain in-person treatment.

Regarding socio-demographic moderators, older age was related to diminished efficacy of guided i-CBT on anxiety and depression but not in targeting other outcomes. This finding concurs with how stronger meta-analytic efficacy was observed for FTF and i-CBT than all controls with working-age adults than their older counterparts on generalized anxiety disorder severity (Kishita & Laidlaw, 2017). Although another meta-analysis reported that the efficacy of FTF CBT wanes for older clients with OCD in the long run (Reid et al., 2021), there were insufficient studies to examine this in the i-CBT context. Future RCTs should investigate how procedural adjustments tailored to older individuals, such as using mnemonics, relevant examples, and simplifying action plans, terms, and instructions, might enhance i-CBT outcomes among older adults. However, as age did not moderate depression, PTSD, and other outcomes in the current and other recent CBT-focused meta-analyses (Paiva et al., 2023; Werson et al., 2022), i-CBT might be a promising scalable treatment across all developmental stages for CMDs.

In addition, it is unclear why self-guided i-CBT had better efficacy on QOL among women than men in the long run. Prior meta-analyses of CBT RCTs found no moderating role of gender on symptom outcomes (Cuijpers et al., 2014; Whiston et al., 2019; Zhang et al., 2019). However, another meta-analysis showed that women were less likely to drop out of self-guided i-CBT than men (Karyotaki et al., 2015), perhaps due to more positive beliefs and engagement with therapy skills (Grubbs et al., 2015).

Notably, there was minimal evidence for dose-response effects (i.e., better treatment efficacy with more exposure or "doses" in terms of i-CBT content or duration defined by weeks of treatment). More i-CBT modules and a greater number of weeks of treatment were associated with greater long-term efficacy of guided/self-guided i-CBT on anxiety severity and QOL. Plausibly, higher "doses" could confer more chances to consolidate CBT skills, such as behavioral experiments and cognitive reframing, thereby alleviating symptoms and enhancing QOL. However, given the primarily non-significant moderator findings, our results largely aligned with prior meta-analytic null dose-response effects in the context of i-CBTs (e.g., Guo et al., 2021). Research suggested that the ideal amount of FTF CBT sessions in routine care varies from 4 to 26 sessions (and 4 to 6 for low-intensity guided self-help; Robinson et al., 2020). Our findings might imply that these patterns generalize to i-CBTs for CMD symptoms and related functional outcomes in the long term.

In addition, the paradox of better clinical outcomes yet lower retention rates when contrasting guided/self-guided i-CBT with all controls is noteworthy. Findings could be explained by how individuals who prefer the ease and flexibility of digital therapies would benefit from engaging with the accessible, modular, and structured approach that either i-CBT offers to manage symptoms and functional outcomes (Walsh & Richards, 2017). However, a lack of personal interactions with a human clinician, technical and technological complexities, and the need for high intrinsic motivation might hinder i-CBT completion and engagement (Andrews et al., 2023). Another explanation of this paradox could be that i-CBT developers and researchers still do not know how much i-CBT is necessary to produce a sufficient effect or improvement. Thus, lack of retention could also mean users no longer felt the need to engage and was not necessarily indicative of treatment non-response. Also, retention rates were comparable in guided i-CBT and FTF CBT, implying that guides (even those not in the user's physical location) could promote engagement and retention.

Our meta-analysis had several limitations. First, most RCTs were conducted in high-income countries (HICs) that hailed from Europe, Australia, and North America, thus constraining generalizability. Future RCTs should test how well i-CBTs sustain gains over long durations in low- and middle-income countries (LMICs) and other continents. Building a robust digital infrastructure through insightful government policies is vital in these efforts to implement, evaluate, and refine i-CBTs in LMICs and other regions and nurture a culture that embeds and embraces digital treatments in healthcare systems (Naderbagi et al., 2024). Such ecosystems are critical to informing cultural adaptations for distinct subpopulations. Second, we did not report racial data, as only U. S. studies tended to report racial composition, wherein most participants identified as White individuals. Customization of i-CBTs for cultural minorities and non-Western contexts may be warranted (Hall et al., 2016; Henrich et al., 2010). Third, crude SES indicators – education and employment status – were used

and found not to moderate i-CBT efficacy. Although this might mean that SES would not modify i-CBT efficacy, thereby signaling the generalizability of i-CBT efficacy across SES classes, future RCTs should consistently administer better SES measures, such as annual household income bracket or housing type. Fourth, older people, perhaps due to a technological disparity, have been inadequately represented. Fifth, many guided i-CBT studies restricted the clinician or coach contact available to users, and some individuals might desire more interaction with their guide. Focusing on the length of clinician time users request instead of the amount delivered might play a crucial moderating role in future research. Sixth, many studies shared similar methodological limitations, such as a lack of blind assessors, neglect to monitor and report adverse events, absence of power calculation, or a missing data management approach. There is a need for more high-quality studies from a methodological standpoint. Seventh, a crucial shortcoming in this literature is the absence of data collected and reported on treatment-seeking behaviors and i-CBT skills usage after the trial intervention phase ends (Puddephatt et al., 2022). Credibility in this area would be increased by capturing and modeling these variables to enhance confidence that any long-term gains observed would be due to practicing i-CBT skills rather than seeking external help. Eighth, it is well-established that CMDs tend to dovetail with high comorbidity rates (Puddephatt et al., 2022). For instance, comorbidities such as personality disorders might attenuate the efficacy of i-CBT (Flygare et al., 2020) but remain under-reported. Future trials should address this knowledge gap. Finally, unlike FTF CBT, i-CBT users often retain access to intervention resources post-program completion, enabling downloadable and easily retrievable i-CBT content, such as portable document files (PDFs). Users are thus empowered to revisit the intervention content anytime, including obtaining screenshots or recording asynchronous interactions with their i-CBT coaches in guided i-CBT. Unfortunately, i-CBT studies with long-term follow-up assessments rarely inquired about or reported these practices. This limitation should be addressed in future long-term trials of i-CBT.

Simultaneously, our meta-analysis had several strengths. Thorough sensitivity analyses were conducted and implied a similar pattern of results. Publication bias analyses suggested that existing biases were unlikely to impact outcomes and interpretations considerably. In addition, given the nature of technology-delivered self-guided i-CBTs and since guided i-CBTs follow standardized protocols, the likelihood of therapist drift, a frequent occurrence in traditional FTF CBT (Speers et al., 2022), is minimized. Finally, we aimed to spur future researchers to include long-term measurements at one-year follow-up and beyond, to help the field better understand how well i-CBTs sustain efficacy gains to target CMD symptoms and related outcomes. Improved knowledge about the promises and pitfalls of the long-term efficacy of low-cost and scalable i-CBTs might contribute pragmatically to triage decisions in stepped-care and stratified resource-limited clinical settings and offer realistic expectations to users with CMD symptoms.

Several clinical implications merit attention. First, at the population level, these findings offer promising evidence that both guided and self-guided i-CBT show comparable effectiveness and superiority compared to TAU. Moreover, less expensive guided/self-guided i-CBT displayed comparable efficacy with more expensive FTF CBT for most outcomes, and i-CBT fared better in targeting anxiety. Thus, promoting i-CBT for CMDs in school

and work organizations may sustainably reach a broad demographic of youths and adults at a presumably reduced cost and with enhanced scalability potential. However, such efforts must be conducted ethically by stating existing limitations from the outset, such as waning efficacy over time and long-term efficacy only for specific CMD symptoms (i.e., anxiety, depression, insomnia) and related outcomes (e.g., remission, ER). Thus, further work investigating the impact of booster sessions on maintaining gains over time could be beneficial to guide treatment optimization efforts. More RCTs are needed to draw firm conclusions about the long-term efficacy of i-CBT for OCD and PTSD symptoms. Second, it is plausible that specific i-CBT components drive the reduction of particular targets in the long run. Future dismantling i-CBT RCTs (e.g., Beukes et al., 2021) should test if, for example, a component (e.g., exposure) worked best to maintain long-term improvements in an outcome (e.g., anxiety). Third, the long-term change mechanisms of i-CBT remain open to inquiry. Future RCTs, using latent change methods, for instance (Alpert et al., 2023), should test how decreases in inflexible thinking or cognitive distortions, which were consistent mediators of FTF CBT (Moreno-Peral et al., 2020; Parsons et al., 2021), might extend to i-CBTs over a year or more. Finally, clinical psychological science would profit from identifying subgroups for whom i-CBTs confer the greatest long-term efficacy to inform precision mental health efforts (Weisz et al., 2023).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding Acknowledgements

Nur Hani Zainal reports financial support was provided by National University of Singapore through the Presidential Young Professorship (PYP) Start-Up Grant. Natalia Van Doren reports financial support was provided by National Institute on Drug Abuse (T32DA007250). Corina Benjet reports financial support was provided by National Institute of Mental Health (1R01MH120648-01A1). None of the funding agencies mentioned here are responsible for the data analysis, writing, and interpretation of the present manuscript. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## References

Adhikary D, Barman S, & Ranjan R (2023). Internet-based cognitive behavioural therapy for individuals with depression and chronic health conditions: A systematic review. Cureus, 15(4), Article e37822. 10.7759/cureus.37822

Alpert E, Shotwell Tabke C, Cole TA, Lee DJ, & Sloan DM (2023). A systematic review of literature examining mediators and mechanisms of change in empirically supported treatments for posttraumatic stress disorder. Clinical Psychology Review, 103, Article 102300. 10.1016/j.cpr.2023.102300

Andersson G (2016). Internet-delivered psychological treatments. Annual Review of Clinical Psychology, 12, 157–179. 10.1146/annurev-clinpsy-021815-093006

Andersson G (2024). The latest developments with internet-based psychological treatments for depression. Expert Review of Neurotherapeutics, 24(2), 171–176. 10.1080/14737175.2024.2309237 [PubMed: 38277244]

Andersson G, Carlbring P, Titov N, & Lindefors N (2019). Internet interventions for adults with anxiety and mood disorders: A narrative umbrella review of recent meta-analyses. Canadian Journal of Psychiatry, 64(7), 465–470. 10.1177/0706743719839381 [PubMed: 31096757]

Andersson G, Rozental A, Shafran R, & Carlbring P (2018). Long-term effects of internet-supported cognitive behaviour therapy. Expert Review of Neurotherapeutics, 18(1), 21–28. 10.1080/14737175.2018.1400381 [PubMed: 29094622]

Andrews B, Klein B, Van Nguyen H, Corboy D, McLaren S, & Watson S (2023). Efficacy of a digital mental health biopsychosocial transdiagnostic intervention with or without therapist assistance for adults with anxiety and depression: Adaptive randomized controlled trial. Journal of Medical Internet Research, 25, Article e45135. 10.2196/45135

Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, & Newby JM (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. Journal of Anxiety Disorders, 55, 70–78. 10.1016/j.janxdis.2018.01.001 [PubMed: 29422409]

Andrews G, & Williams AD (2015). Up-scaling clinician assisted internet cognitive behavioural therapy (iCBT) for depression: A model for dissemination into primary care. Clinical Psychology Review, 41, 40–48. 10.1016/j.cpr.2014.05.006 [PubMed: 25043445]

Bai N, Cao J, Zhang H, Liu X, & Yin M (2024). Digital cognitive behavioural therapy for patients with insomnia and depression: A systematic review and meta-analysis. Journal of Psychiatric and Mental Health Nursing, 31(4), 654–667. 10.1111/jpm.13024 [PubMed: 38226714]

Bai P (2023). Application and mechanisms of internet-based cognitive behavioral therapy (iCBT) in improving psychological state in cancer patients. Journal of Cancer, 14(11), 1981–2000. 10.7150/jca.82632 [PubMed: 37497400]

van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, & Riper H (2014). Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: A meta-analysis. PLoS One, 9(7), Article e100674. 10.1371/journal.pone.0100674

Bandelow B, Allgulander C, Baldwin DS, Costa DLDC, Denys D, Dilbaz N, & Zohar J (2023). World Federation of Societies of biological psychiatry (WFSBP) guidelines for treatment of anxiety, obsessive-compulsive and posttraumatic stress disorders – Version 3. Part I: Anxiety disorders. The World Journal of Biological Psychiatry, 24(2), 79–117. 10.1080/15622975.2022.2086295 [PubMed: 35900161]

Baumeister H, Reichler L, Munzinger M, & Lin J (2014). The impact of guidance on internet-based mental health interventions — A systematic review. Internet Interventions, 1(4), 205–215. 10.1016/j.invent.2014.08.003

Baxter AJ, Vos T, Scott KM, Ferrari AJ, & Whiteford HA (2014). The global burden of anxiety disorders in 2010. Psychological Medicine, 44(11), 2363–2374. 10.1017/S0033291713003243 [PubMed: 24451993]

Beatty L, & Binnion C (2016). A systematic review of predictors of, and reasons for, adherence to online psychological interventions. International Journal of Behavioral Medicine, 23(6), 776–794. 10.1007/s12529-016-9556-9 [PubMed: 26957109]

Beck AT, & Haigh EAP (2014). Advances in cognitive theory and therapy: The generic cognitive model. Annual Review of Clinical Psychology, 10(1), 1–24. 10.1146/annurev-clinpsy-032813-153734

Begg CB, & Mazumdar M (1994). Operating characteristics of a rank correlation test for publication bias. Biometrics, 50(4), 1088–1101. [PubMed: 7786990]

Beintner I, Vollert B, Zarski AC, Bolinski F, Musiat P, Gorlich D, & Jacobi C (2019). Adherence reporting in randomized controlled trials examining manualized multisession online interventions: Systematic review of practices and proposal for reporting standards. Journal of Medical Internet Research, 21(8), Article e14181. 10.2196/14181

Beukes EW, Andersson G, Fagelson MA, & Manchaiah V (2021). Dismantling internet-based cognitive behavioral therapy for tinnitus. The contribution of applied relaxation: A randomized controlled trial. Internet Interventions, 25, Article 100402. 10.1016/j.invent.2021.100402

Borenstein M, Hedges LV, Higgins JPT, & Rothstein HR (2009). Introduction to meta-analysis. John Wiley & Sons. 10.1002/9780470743386

Brantnell A, Woodford J, Baraldi E, van Achterberg T, & von Essen L (2020). Views of implementers and nonimplementers of internet-administered cognitive behavioral therapy for depression and anxiety: Survey of primary care decision makers in Sweden. Journal of Medical Internet Research, 22(8), Article e18033. 10.2196/18033

Cheung MW (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. Psychological Methods, 13(3), 182–202. 10.1037/a0013163 [PubMed: 18778151]

Cheung MW (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. Psychological Methods, 19(2), 211–229. 10.1037/a0032968 [PubMed: 23834422]

Clark DM (2018). Realizing the mass public benefit of evidence-based psychological therapies: The IAPT program. Annual Review of Clinical Psychology, 14, 159–183. 10.1146/annurev-clinpsy-050817-084833

Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, & Thornicroft G (2015). What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychological Medicine, 45(1), 11–27. 10.1017/S0033291714000129 [PubMed: 24569086]

Cochran WG (1954). The combination of estimates from different experiments. Biometrics, 10(1), 101–129. 10.2307/3001666

Cohen J (1988). Statistical power for the social sciences. Hillsdale, NJ: Laurence Erlbaum and Associates.

Cuijpers P, Miguel C, Harrer M, Plessen CY, Ciharova M, Ebert D, & Karyotaki E (2023). Cognitive behavior therapy vs. control conditions, other psychotherapies, pharmacotherapies and combined treatment for depression: A comprehensive meta-analysis including 409 trials with 52,702 patients. World Psychiatry, 22(1), 105–115. 10.1002/wps.21069 [PubMed: 36640411]

Cuijpers P, Weitz E, Twisk J, Kuehner C, Cristea I, David D, & Hollon SD (2014). Gender as predictor and moderator of outcome in cognitive behavior therapy and pharmacotherapy for adult depression: An “individual patient data” meta-analysis. Depression and Anxiety, 31(11), 941–951. 10.1002/da.22328 [PubMed: 25407584]

Cusack K, Jonas DE, Forneris CA, Wines C, Sonis J, Middleton JC, & Gaynes BN (2016). Psychological treatments for adults with posttraumatic stress disorder: A systematic review and meta-analysis. Clinical Psychology Review, 43, 128–141. 10.1016/j.cpr.2015.10.003 [PubMed: 26574151]

Dalle Grave R, Sartirana M, & Calugi S (2020). Personalized cognitive-behavioural therapy for obesity (CBT-OB): Theory, strategies and procedures. BioPsychoSocial Medicine, 14, 5. 10.1186/s13030-020-00177-9 [PubMed: 32175002]

Daros AR, Haefner SA, Asadi S, Kazi S, Rodak T, & Quilty LC (2021). A meta-analysis of emotional regulation outcomes in psychological interventions for youth with depression and anxiety. Nature Human Behaviour, 5(10), 1443–1457. 10.1038/s41562-021-01191-9

Deeks JJ, Higgins JPT, & Altman DG (2023). Chapter 10: Analysing data and undertaking meta-analyses. In Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, & Welch VA (Eds.), Cochrane handbook for systematic reviews of interventions version 6.4 (updated august 2023). Cochrane www.training.cochrane.org/handbook.

Dulsen P, & Baumeister H (2024). Internet- and mobile-based anxiety and depression interventions for children and adolescents: Efficacy and negative effects - a systematic review and meta-analysis. European Child & Adolescent Psychiatry.. 10.1007/s00787-024-02404-y

Duval S, & Tweedie R (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56(2), 455–463. 10.1111/j.0006-341x.2000.00455.x [PubMed: 10877304]

Eilert N, Wogan R, Adegoke A, Earley C, Duffy D, Enrique A, & Richards D (2023). The relationship between posttherapeutic cognitive behavior therapy skills usage and follow-up outcomes of internet-delivered cognitive behavior therapy. Journal of Clinical Psychology, 79(1), 55–67. 10.1002/jclp.23403 [PubMed: 35726497]

Fawcett EJ, Power H, & Fawcett JM (2020). Women are at greater risk of OCD than men: A meta-analytic review of OCD prevalence worldwide. Journal of Clinical Psychiatry, 81(4). 10.4088/JCP.19r13085

Fernandez-Rodriguez R, Zhao L, Bizzozero-Peroni B, Martinez-Vizcaino V, Mesas AE, Wittert G, & Heilbronn LK (2024). Are e-health interventions effective in reducing diabetes-related distress and depression in patients with type 2 diabetes? A systematic review with meta-analysis. Telemedicine Journal and E-Health, 30(4), 919–939. 10.1089/tmj.2023.0374 [PubMed: 38010739]

Fitzsimmons-Craft EE, Rojas E, Topooco N, Rackoff GN, Zainal NH, Eisenberg D, & Newman MG (2023). Training, supervision, and experience of coaches offering digital guided self-help for mental health concerns. Frontiers in Psychology, 14, 1–9. 10.3389/fpsyg.2023.1217698

Flygare AL, Engstrom I, Hasselgren M, Jansson-Frojmark M, Frejgrim R, Andersson G, & Hollandare F (2020). Internet-based CBT for patients with depressive disorders in primary and psychiatric care: Is it effective and does comorbidity affect outcome? Internet Interventions, 19, Article 100303. 10.1016/j.invent.2019.100303

Georgieva I, Lepping P, Bozev V, Lickiewicz J, Pekara J, Wikman S, & Lantta T (2021). Prevalence, new incidence, course, and risk factors of PTSD, depression, anxiety, and panic disorder during the Covid-19 pandemic in 11 countries. Healthcare, 9(6). 10.3390/healthcare9060664

Goetter EM, Frumkin MR, Palitz SA, Swee MB, Baker AW, Bui E, & Simon NM (2020). Barriers to mental health treatment among individuals with social anxiety disorder and generalized anxiety disorder. Psychological Services, 17 (1), 5–12. 10.1037/ser0000254

Grubbs KM, Cheney AM, Fortney JC, Edlund C, Han X, Dubbert P, & Sullivan JG (2015). The role of gender in moderating treatment outcome in collaborative care for anxiety. Psychiatric Services, 66(3), 265–271. 10.1176/appi.ps.201400049 [PubMed: 25727114]

Guo S, Deng W, Wang H, Liu J, Liu X, Yang X, & Li X (2021). The efficacy of internet-based cognitive behavioural therapy for social anxiety disorder: A systematic review and meta-analysis. Clinical Psychology & Psychotherapy, 28(3), 656–668. 10.1002/cpp.2528 [PubMed: 33142004]

Gutierrez G, Gizzarelli T, Moghimi E, Vazquez G, & Alavi N (2023). Online cognitive behavioral therapy (eCBT) for the management of depression symptoms in unipolar and bipolar spectrum disorders, a systematic review and network meta-analysis. Journal of Affective Disorders, 341, 379–392. 10.1016/j.jad.2023.09.005 [PubMed: 37683940]

Hadjistavropoulos HD, Peynenburg V, Thiessen DL, Nugent M, Karin E, Dear BF, & Titov N (2022). A randomized factorial trial of internet-delivered cognitive behavioural therapy: An 8-week program with or without extended support and booster lesson. Internet Interventions, 27, Article 100499. 10.1016/j.invent.2022.100499

Hall GC, Ibaraki AY, Huang ER, Marti CN, & Stice E (2016). A meta-analysis of cultural adaptations of psychological interventions. Behavior Therapy, 47(6), 993–1014. 10.1016/j.beth.2016.09.005 [PubMed: 27993346]

Hasanvandi S, Bakhtiari M, Askarirostami F, & Ghiasi H (2022). Meta-analysis of the effectiveness of online cognitive-behavioral therapy in the treatment of post-traumatic stress disorder. Medical Science, 26(122). 10.54905/disssi/v26i122/ms119e2012

Hedges LV, Tipton E, & Johnson MC (2010). Robust variance estimation in meta-regression with dependent effect size estimates. Research Synthesis Methods, 1(1), 39–65. 10.1002/jrsm.5 [PubMed: 26056092]

Henrich J, Heine SJ, & Norenzayan A (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33(2–3), 61–83. 10.1017/S0140525X0999152X [PubMed: 20550733]

Higgins JP, & Thompson SG (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539–1558. 10.1002/sim.1186 [PubMed: 12111919]

Hoo KA, Tvarlapati KJ, Piovoso MJ, & Hajare R (2002). A method of robust multivariate outlier replacement. Computers & Chemical Engineering, 26(1), 17–39. 10.1016/S0098-1354(01)00734-7

Johnson E, Corrick S, Isley S, Vandermeer B, Dolgoy N, Bates J, & Tandon P (2024). Mind-body internet and mobile-based interventions for depression and anxiety in adults with chronic physical conditions: A systematic review of RCTs. PLOS Digit Health, 3(1), Article e0000435. 10.1371/journal.pdig.0000435

Jonsson U, Linton SJ, Ybrandt H, Ringborg A, Leander L, Moberg K, & Arnberg FK (2023). Internet-delivered psychological treatment as an add-on to treatment as usual for common mental disorders: A systematic review with meta-analysis of randomized trials. Journal of Affective Disorders, 322, 221–234. 10.1016/j.jad.2022.11.036 [PubMed: 36400149]

Kambach S, Bruelheide H, Gerstner K, Gurevitch J, Beckmann M, & Seppelt R (2020). Consequences of multiple imputation of missing standard deviations and sample sizes in meta-analysis. Ecology and Evolution, 10(20), 11699–11712. 10.1002/ece3.6806 [PubMed: 33144994]

Karyotaki E, Efthimiou O, Miguel C, Bermpohl FMG, Furukawa TA, Cuijpers P, & Forsell Y (2021). Internet-based cognitive behavioral therapy for depression: A systematic review and individual patient data network meta-analysis. JAMA Psychiatry, 78(4), 361–371. 10.1001/jamapsychiatry.2020.4364 [PubMed: 33471111]

Karyotaki E, Kemmeren L, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, & Cuijpers P (2018). Is self-guided internet-based cognitive behavioural therapy (iCBT) harmful? An individual participant data meta-analysis. Psychological Medicine, 48(15), 2456–2466. 10.1017/S0033291718000648 [PubMed: 29540243]

Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, & Cuijpers P (2015). Predictors of treatment dropout in self-guided web-based interventions for depression: An “individual patient data” meta-analysis. Psychological Medicine, 45(13), 2717–2726. 10.1017/S0033291715000665 [PubMed: 25881626]

Kishita N, & Laidlaw K (2017). Cognitive behaviour therapy for generalized anxiety disorder: Is CBT equally efficacious in adults of working age and older adults? Clinical Psychology Review, 52, 124–136. 10.1016/j.cpr.2017.01.003 [PubMed: 28119196]

Kohnen M, Kriston L, Harter M, Baumeister H, & Liebherz S (2021). Effectiveness and acceptance of technology-based psychological interventions for the acute treatment of unipolar depression: Systematic review and meta-analysis. Journal of Medical Internet Research, 23(6), Article e24584. 10.2196/24584

Koo TK, & Li MY (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155–163. 10.1016/j.jcm.2016.02.012 [PubMed: 27330520]

Krzyzaniak N, Greenwood H, Scott AM, Peiris R, Cardona M, Clark J, & Glasziou P (2024). The effectiveness of telehealth versus face-to face interventions for anxiety disorders: A systematic review and meta-analysis. Journal of Telemedicine and Telecare, 30(2), 250–261. 10.1177/1357633X211053738 [PubMed: 34860613]

Kuester A, Niemeyer H, & Knaevelsrud C (2016). Internet-based interventions for posttraumatic stress: A meta-analysis of randomized controlled trials. Clinical Psychology Review, 43, 1–16. 10.1016/j.cpr.2015.11.004 [PubMed: 26655959]

Kuhne F, Ay DS, Marschner L, & Weck F (2020). The heterogeneous course of OCD - a scoping review on the variety of definitions. Psychiatry Research, 285, Article 112821. 10.1016/j.psychres.2020.112821

Kumar V, Sattar Y, Bseiso A, Khan S, & Rutkofsky IH (2017). The effectiveness of internet-based cognitive behavioral therapy in treatment of psychiatric disorders. Cureus, 9(8), Article e1626. 10.7759/cureus.1626

Leichsenring F, Steinert C, Rost F, Abbass A, Heim N, & Ioannidis JPA (2023). A critical assessment of NICE guidelines for treatment of depression. World Psychiatry, 22(1), 43–45. 10.1002/wps.21039 [PubMed: 36640399]

Lewis C, Roberts NP, Simon N, Bethell A, & Bisson JI (2019). Internet-delivered cognitive behavioural therapy for post-traumatic stress disorder: Systematic review and meta-analysis. Acta Psychiatrica Scandinavica, 140(6), 508–521. 10.1111/acps.13079 [PubMed: 31359407]

Linardon J, Torous J, Firth J, Cuijpers P, Messer M, & Fuller-Tyszkiewicz M (2024). Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry, 23(1), 139–149. 10.1002/wps.21183 [PubMed: 38214614]

Lo CKM, Chan KL, Yu L, Chui WWH, & Ip P (2023). Long-term effects of psychosocial interventions on internet-related disorders: A meta-analysis. Computers in Human Behavior, 138. 10.1016/j.chb.2022.107465

Lo HK, Ho FY, Yeung JW, Ng ST, Wong EY, & Chung KF (2024). Self-help interventions for the prevention of relapse in mood disorder: A systematic review and meta-analysis. Family Practice. 10.1093/fampra/cmae036

Luo C, Sanger N, Singhal N, Pattrick K, Shams I, Shahid H, & Samaan Z (2020). A comparison of electronically-delivered and face to face cognitive behavioural therapies in depressive disorders: A systematic review and meta-analysis. EClinicalMedicine, 24, Article 100442. 10.1016/j.eclinm.2020.100442

Maj A, Michalak N, Graczykowska A, & Andersson G (2023). The effect of internet-delivered cognitive behavioral therapy for depression and anxiety on quality of life: A meta-analysis of randomized controlled trials. Internet Interventions, 33, Article 100654. 10.1016/j.invent.2023.100654

Mamukashvili-Delau M, Koburger N, Dietrich S, & Rummel-Kluge C (2023). Long-term efficacy of internet-based cognitive behavioral therapy self-help programs for adults with depression: Systematic review and meta-analysis of randomized controlled trials. JMIR Mental Health, 10, Article e46925. 10.2196/46925

Mayer M (2024). missRanger: Fast imputation of missing values. In Version 2.5.0. https://CRAN.R-project.org/package=missRanger.

McCall H, Dear BF, Landry C, Beahm JD, Gregory J, Titov N, & Hadjistavropoulos HD (2023). Internet-delivered cognitive behavioural therapy for symptoms of PTSD among public safety personnel: Initial outcomes of an open cohort preference trial of transdiagnostic and disorder-specific therapy. Internet Interventions, 33, Article 100656. 10.1016/j.invent.2023.100656

McGuinness LA (2019). Robvis: An R package and web application for visualising risk-of-bias assessments. https://github.com/mcguinlu/robvis.

McHugh RK, Whitton SW, Peckham AD, Welge JA, & Otto MW (2013). Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: A meta-analytic review. Journal of Clinical Psychiatry, 74(6), 595–602. 10.4088/JCP.12r07757 [PubMed: 23842011]

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, & Group P-P (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews, 4(1), 1. 10.1186/2046-4053-4-1 [PubMed: 25554246]

Moltrecht B, Deighton J, Patalay P, & Edbrooke-Childs J (2021). Effectiveness of current psychological interventions to improve emotion regulation in youth: A meta-analysis. European Child & Adolescent Psychiatry, 30(6), 829–848. 10.1007/s00787-020-01498-4 [PubMed: 32108914]

Moreno-Peral P, Bellon JA, Huibers MJH, Mestre JM, Garcia-Lopez LJ, Taubner S, & Conejo-Ceron S (2020). Mediators in psychological and psychoeducational interventions for the prevention of depression and anxiety. A systematic review. Clinical Psychology Review, 76, Article 101813. 10.1016/j.cpr.2020.101813

Morris SB, & DeShon RP (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological Methods, 7, 105–125. 10.1037/1082-989X.7.1.105 [PubMed: 11928886]

Moshe I, Terhorst Y, Philippi P, Domhardt M, Cuijpers P, Cristea I, & Sander LB (2021). Digital interventions for the treatment of depression: A meta-analytic review. Psychological Bulletin, 147(8), 749–786. 10.1037/bul0000334 [PubMed: 34898233]

Naderbagi A, Loblay V, Zahed IUM, Ekambareshwar M, Poulsen A, Song YJC, & LaMonica HM (2024). Cultural and contextual adaptation of digital health interventions: Narrative review. Journal of Medical Internet Research, 26, Article e55130. 10.2196/55130

Nakao M, Shirotsuki K, & Sugaya N (2021). Cognitive-behavioral therapy for management of mental health and stress-related disorders: Recent advances in techniques and technologies. BioPsychoSocial Medicine, 15(1), 16. 10.1186/s13030-021-00219-w [PubMed: 34602086]

Oey LT, McDonald S, McGrath L, Dear BF, & Wootton BM (2023). Guided versus self-guided internet delivered cognitive behavioural therapy for diagnosed anxiety and related disorders: A preliminary meta-analysis. Cognitive Behaviour Therapy, 52(6), 654–671. 10.1080/16506073.2023.2250073 [PubMed: 37655553]

Paiva JM, Dos Santos Melani M, Marques ESN, Arcosy CV, Coutinho ESF, Ventura P, & Berger W (2023). The efficacy of internet-delivered cognitive-behavioral therapy for posttraumatic stress disorder according to the mean age of patients: A systematic review and meta-analysis. Psychology, Health & Medicine, 1–15. 10.1080/13548506.2023.2292478

Parsons CE, Zachariae R, Landberger C, & Young KS (2021). How does cognitive behavioural therapy for insomnia work? A systematic review and meta-analysis of mediators of change. Clinical Psychology Review, 86, Article 102027. 10.1016/j.cpr.2021.102027

Pauley D, Cuijpers P, Papola D, Miguel C, & Karyotaki E (2021). Two decades of digital interventions for anxiety disorders: A systematic review and meta-analysis of treatment effectiveness. Psychological Medicine, 1–13. 10.1017/S0033291721001999

Peynenburg V, Hadjistavropoulos H, Thiessen D, Titov N, & Dear B (2022). Internet-delivered cognitive behavioral therapy for postsecondary students: Randomized factorial trial for examining motivational interviewing and booster lessons. Journal of Medical Internet Research, 24(9), Article e40637. 10.2196/40637

Polak M, & Tanzer NK (2024). Internet-based cognitive behavioural treatments for obsessive-compulsive disorder: A systematic review and meta-analysis. Clinical Psychology & Psychotherapy, 31(3), Article e2989. 10.1002/cpp.2989

Puddephatt JA, Irizar P, Jones A, Gage SH, & Goodwin L (2022). Associations of common mental disorder with alcohol use in the adult general population: A systematic review and meta-analysis. Addiction, 117(6), 1543–1572. 10.1111/add.15735 [PubMed: 34729837]

R Core Team. (2024). R: A language and environment for statistical computing. In Version 4.4.1 R foundation for statistical computing. https://www.R-project.org/.

Reid JE, Laws KR, Drummond L, Vismara M, Grancini B, Mpavaenda D, & Fineberg NA (2021). Cognitive behavioural therapy with exposure and response prevention in the treatment of obsessive-compulsive disorder: A systematic review and meta-analysis of randomised controlled trials. Comprehensive Psychiatry, 106, Article 152223. 10.1016/j.comppsych.2021.152223

Revelle W (2024). Psych: Procedures for psychological, psychometric, and personality research. In Version R package version 2.4.6. Northwestern University. https://CRAN.R-project.org/package=psych.

Robinson L, Delgadillo J, & Kellett S (2020). The dose-response effect in routinely delivered psychological therapies: A systematic review. Psychotherapy Research, 30 (1), 79–96. 10.1080/10503307.2019.1566676 [PubMed: 30661486]

Saether SMM, Knapstad M, Grey N, & Smith ORF (2019). Twelve months post-treatment results from the Norwegian version of improving access to psychological therapies. Frontiers in Psychology, 10, 2303. 10.3389/fpsyg.2019.02303 [PubMed: 31681099]

Santomauro DF, Mantilla Herrera AM, Shadid J, Zheng P, Ashbaugh C, & COVID-19 Mental Disorders Collaborators. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet, 398(10312), 1700–1712. 10.1016/S0140-6736(21)02143-7 [PubMed: 34634250]

Schure M, McCrory B, Tuchscherer Franklin K, Greist J, & Weissman RS (2020). Twelve-month follow-up to a fully automated internet-based cognitive behavior therapy intervention for rural adults with depression symptoms: Single-arm longitudinal study. Journal of Medical Internet Research, 22(10), Article e21336. 10.2196/21336

Simes RJ (1986). An improved Bonferroni procedure for multiple tests of significance. Biometrika, 73(3), 751–754. 10.2307/2336545

Simon L, Steinmetz L, Feige B, Benz F, Spiegelhalder K, & Baumeister H (2023). Comparative efficacy of onsite, digital, and other settings for cognitive behavioral therapy for insomnia: A systematic review and network meta-analysis. Scientific Reports, 13(1), 1929. 10.1038/s41598-023-28853-0 [PubMed: 36732610]

Speers AJH, Bhullar N, Cosh S, & Wootton BM (2022). Correlates of therapist drift in psychological practice: A systematic review of therapist characteristics. Clinical Psychology Review, 93, Article 102132. 10.1016/j.cpr.2022.102132

Springer KS, Levy HC, & Tolin DF (2018). Remission in CBT for adult anxiety disorders: A meta-analysis. Clinical Psychology Review, 61, 1–8. 10.1016/j.cpr.2018.03.002 [PubMed: 29576326]

Steinert C, Hofmann M, Leichsenring F, & Kruse J (2015). The course of PTSD in naturalistic long-term studies: High variability of outcomes. A systematic review. Nordic Journal of Psychiatry, 69(7), 483–496. 10.3109/08039488.2015.1005023 [PubMed: 25733025]

Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, & Higgins JPT (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. BMJ, 366, Article l4898. 10.1136/bmj.l4898

Struijs SY, Lamers F, Rinck M, Roelofs K, Spinhoven P, & Penninx B (2018). The predictive value of approach and avoidance tendencies on the onset and course of depression and anxiety disorders. Depression and Anxiety, 35(6), 551–559. 10.1002/da.22760 [PubMed: 29689597]

Svardman F, Sjowall D, & Lindsater E (2022). Internet-delivered cognitive behavioral interventions to reduce elevated stress: A systematic review and meta-analysis. Internet Interventions, 29, Article 100553. 10.1016/j.invent.2022.100553

Sztein DM, Koransky CE, Fegan L, & Himelhoch S (2018). Efficacy of cognitive behavioural therapy delivered over the internet for depressive symptoms: A systematic review and meta-analysis. Journal of Telemedicine and Telecare, 24(8), 527–539. 10.1177/1357633X17717402 [PubMed: 28696153]

Terhorst Y, Kaiser T, Brakemeier EL, Moshe I, Philippi P, Cuijpers P, & Sander LB (2024). Heterogeneity of treatment effects in internet- and mobile-based interventions for depression: A systematic review and meta-analysis. JAMA Network Open, 7(7), Article e2423241. 10.1001/jamanetworkopen.2024.23241

Tipton E (2013). Robust variance estimation in meta-regression with binary dependent effects. Research Synthesis Methods, 4(2), 169–187. 10.1002/jrsm.1070 [PubMed: 26053656]

Tipton E, & Pustejovsky JE (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. Journal of Educational and Behavioral Statistics, 40(6), 604–634. 10.3102/1076998615606099

Tng GYQ, Koh J, Soh XC, Majeed NM, & Hartanto A (2024). Efficacy of digital mental health interventions for PTSD symptoms: A systematic review of meta-analyses. Journal of Affective Disorders, 357, 23–36. 10.1016/j.jad.2024.04.074 [PubMed: 38679205]

Tolin DF (2016). Doing CBT: A comprehensive guide to working with behaviors, thoughts, and emotions. The Guilford Press.

Tong AC, Ho FS, Chu OH, & Mak WW (2023). Time-dependent changes in depressive symptoms among control participants in digital-based psychological intervention studies: Meta-analysis of randomized controlled trials. Journal of Medical Internet Research, 25, Article e39029. 10.2196/39029

Tong L, Panagiotopoulou OM, Cuijpers P, & Karyotaki E (2024). The effectiveness of self-guided interventions in adults with depressive symptoms: A systematic review and meta-analysis. EBioMedicine, 105, Article 105208. 10.1016/j.ebiom.2024.105208

Viechtbauer W (2010). Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36(3), 1–48. 10.18637/jss.v036.i03

Viechtbauer W, & Cheung MW-L (2010). Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods, 1(2), 112–125. 10.1002/jrsm.11 [PubMed: 26061377]

Vittengl JR, Clark LA, Dunn TW, & Jarrett RB (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioral therapy's effects. Journal of Consulting and Clinical Psychology, 75(3), 475–488. 10.1037/0022-006X.75.3.475 [PubMed: 17563164]

Walsh A, & Richards D (2017). Experiences and engagement with the design features and strategies of an internet-delivered treatment programme for generalised anxiety disorder: A service-based evaluation. British Journal of Guidance & Counselling, 45(1), 16–31. 10.1080/03069885.2016.1153039

Wang M, Chen H, Yang F, Xu X, & Li J (2023). Effects of digital psychotherapy for depression and anxiety: A systematic review and Bayesian network meta-analysis. Journal of Affective Disorders, 338, 569–580. 10.1016/j.jad.2023.06.057 [PubMed: 37392941]

Wasil AR, Palermo EH, Lorenzo-Luaces L, & DeRubeis RJ (2022). Is there an app for that? A review of popular apps for depression, anxiety, and well-being. Cognitive and Behavioral Practice, 29(4), 883–901. 10.1016/j.cbpra.2021.07.001

Weisz JR, Venturo-Conerly KE, Fitzpatrick OM, Frederick JA, & Ng MY (2023). What four decades of meta-analysis have taught us about youth psychotherapy and the science of research synthesis. Annual Review of Clinical Psychology, 19, 79–105. 10.1146/annurev-clinpsy-080921-082920

Werson AD, Meiser-Stedman R, & Laidlaw K (2022). A meta-analysis of CBT efficacy for depression comparing adults and older adults. Journal of Affective Disorders, 319, 189–201. 10.1016/j.jad.2022.09.020 [PubMed: 36113691]

Whiston A, Bockting CLH, & Semkovska M (2019). Towards personalising treatment: A systematic review and meta-analysis of face-to-face efficacy moderators of cognitive-behavioral therapy and interpersonal psychotherapy for major depressive disorder. Psychological Medicine, 49(16), 2657–2668. 10.1017/S0033291719002812 [PubMed: 31615583]

Wickham H, François R, Henry L, Müller K, & Vaughan D (2023). Dpelyr: A grammar of data manipulation. In Version R package version 1.1.4. https://CRAN.R-project.org/package=dplyr.

Wijesuriya R, Moreno-Betancur M, Carlin JB, & Lee KJ (2020). Evaluation of approaches for multiple imputation of three-level data. BMC Medical Research Methodology, 20(1), 207. 10.1186/s12874-020-01079-8 [PubMed: 32787781]

Xiong J, Wen JL, Pei GS, Han X, & He DQ (2023). Effectiveness of internet-based cognitive behavioural therapy for employees with depression: A systematic review and meta-analysis. International Journal of Occupational Safety and Ergonomics, 29(1), 268–281. 10.1080/10803548.2022.2043647 [PubMed: 35172706]

Zainal NH, Chan WW, Saxena AP, Taylor CB, & Newman MG (2021). Pilot randomized trial of self-guided virtual reality exposure therapy for social anxiety disorder. Behaviour Research and Therapy, 147, Article 103984. 10.1016/j.brat.2021.103984

Zainal NH, & Newman MG (2019). Relation between cognitive and behavioral strategies and future change in common mental health problems across 18 years. Journal of Abnormal Psychology, 128(4), 295–304. 10.1037/abn0000428 [PubMed: 31045412]

Zhang A, Borhneimer LA, Weaver A, Franklin C, Hai AH, Guz S, & Shen L (2019). Cognitive behavioral therapy for primary care depression and anxiety: A secondary meta-analytic review using robust variance estimation in meta-regression. Journal of Behavioral Medicine, 42(6), 1117–1141. 10.1007/s10865-019-00046-z [PubMed: 31004323]

Zhang W, Yang W, Ruan H, Gao J, & Wang Z (2023). Comparison of internet-based and face-to-face cognitive behavioral therapy for obsessive-compulsive disorder: A systematic review and network meta-analysis. Journal of Psychiatric Research, 168, 140–148. 10.1016/j.jpsychires.2023.10.025 [PubMed: 37907037]

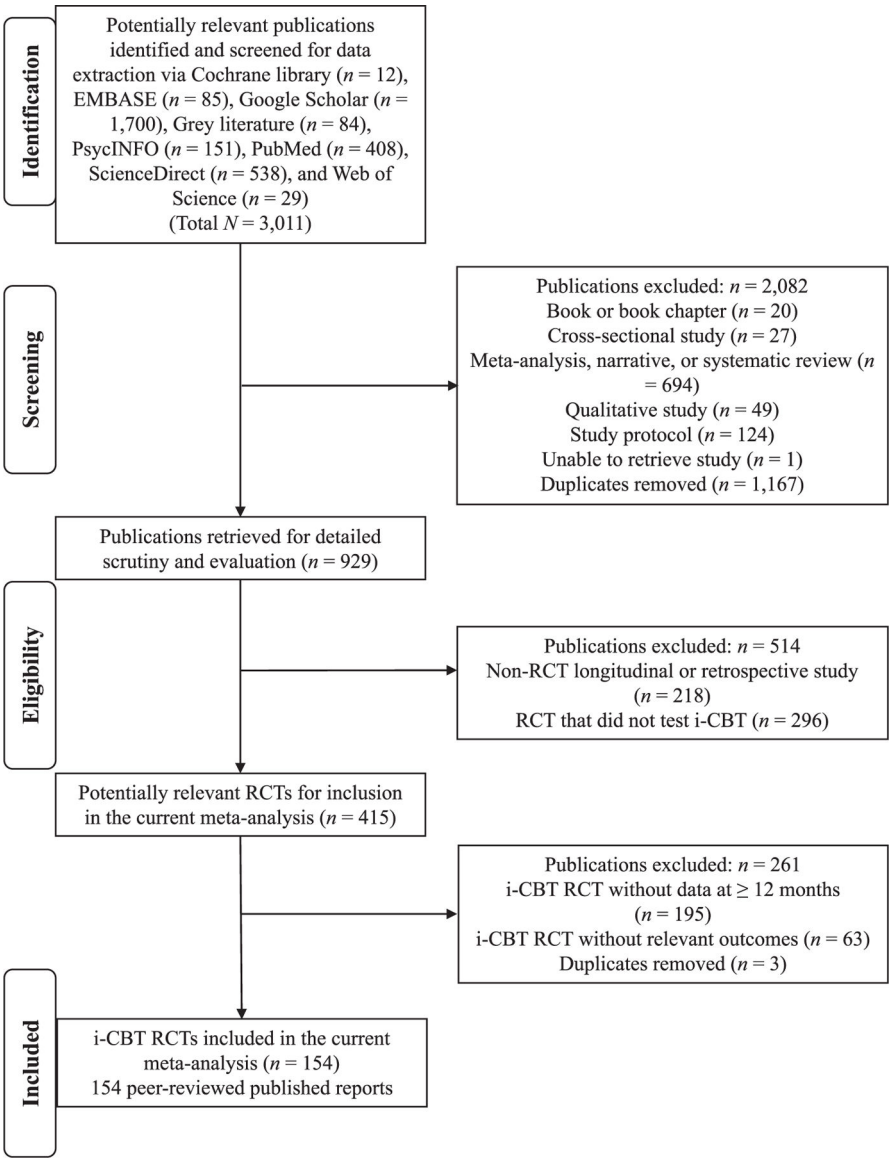


**Fig. 1.**

PRISMA flowchart PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomized controlled trial; i-CBT, internet-delivered cognitive-behavioral therapy.

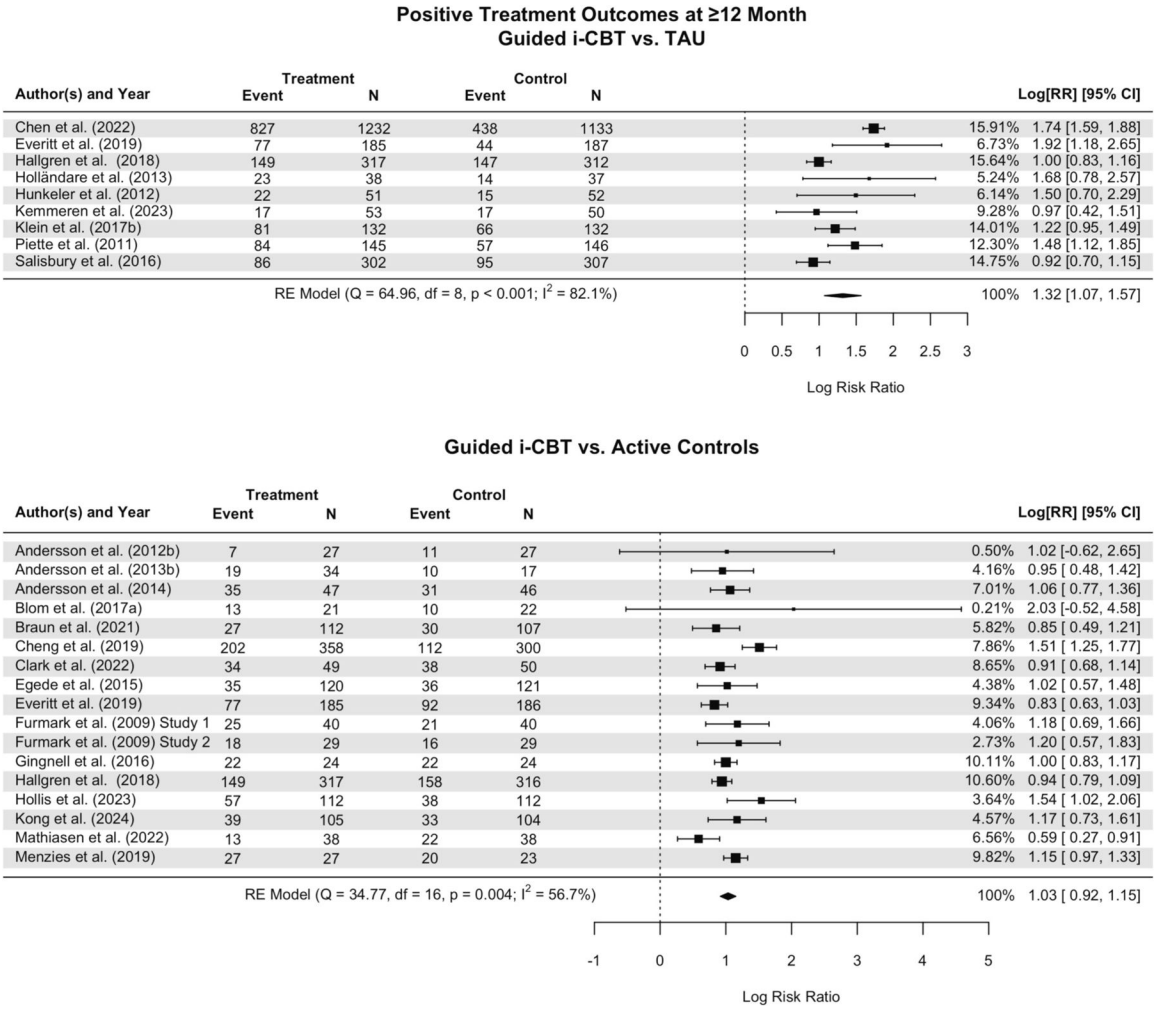


**Fig. 2.**

Long-term efficacy of i-CBT on positive treatment outcomes i-CBT, internet-delivered cognitive-behavioral therapy.

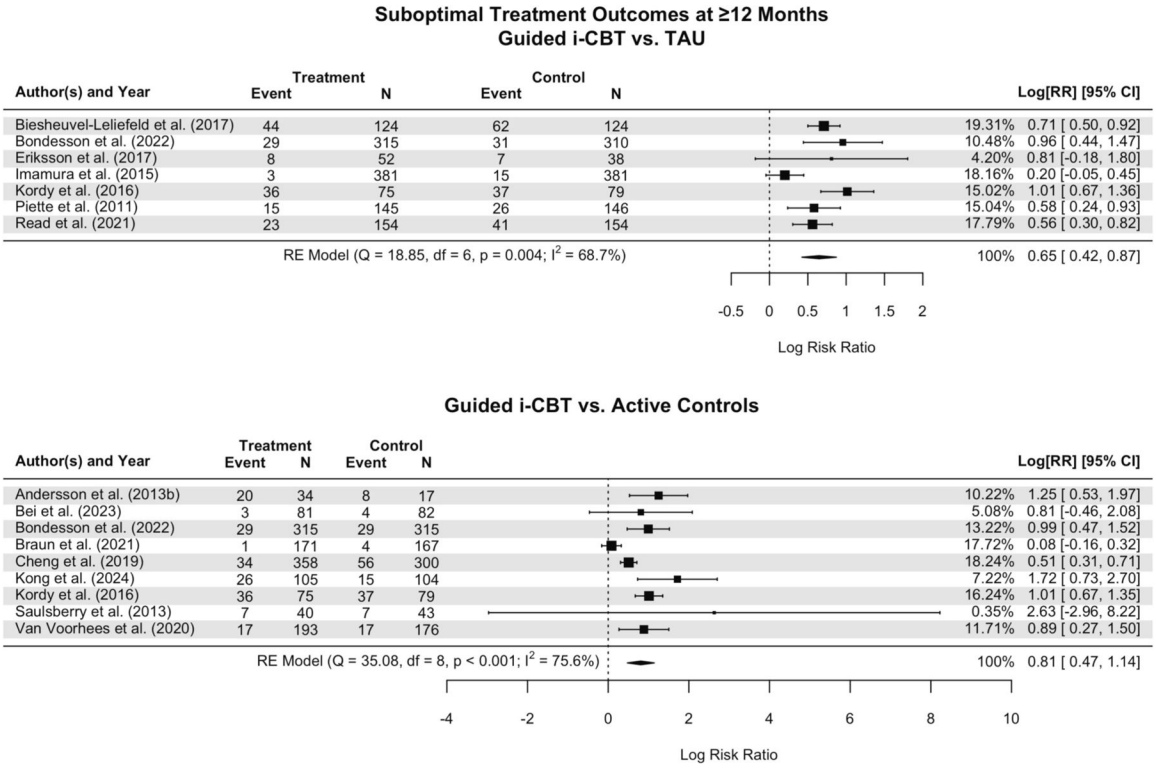


**Fig. 3.**

Long-term efficacy of i-CBT on suboptimal treatment outcomes i-CBT, internet-delivered cognitive-behavioral therapy.

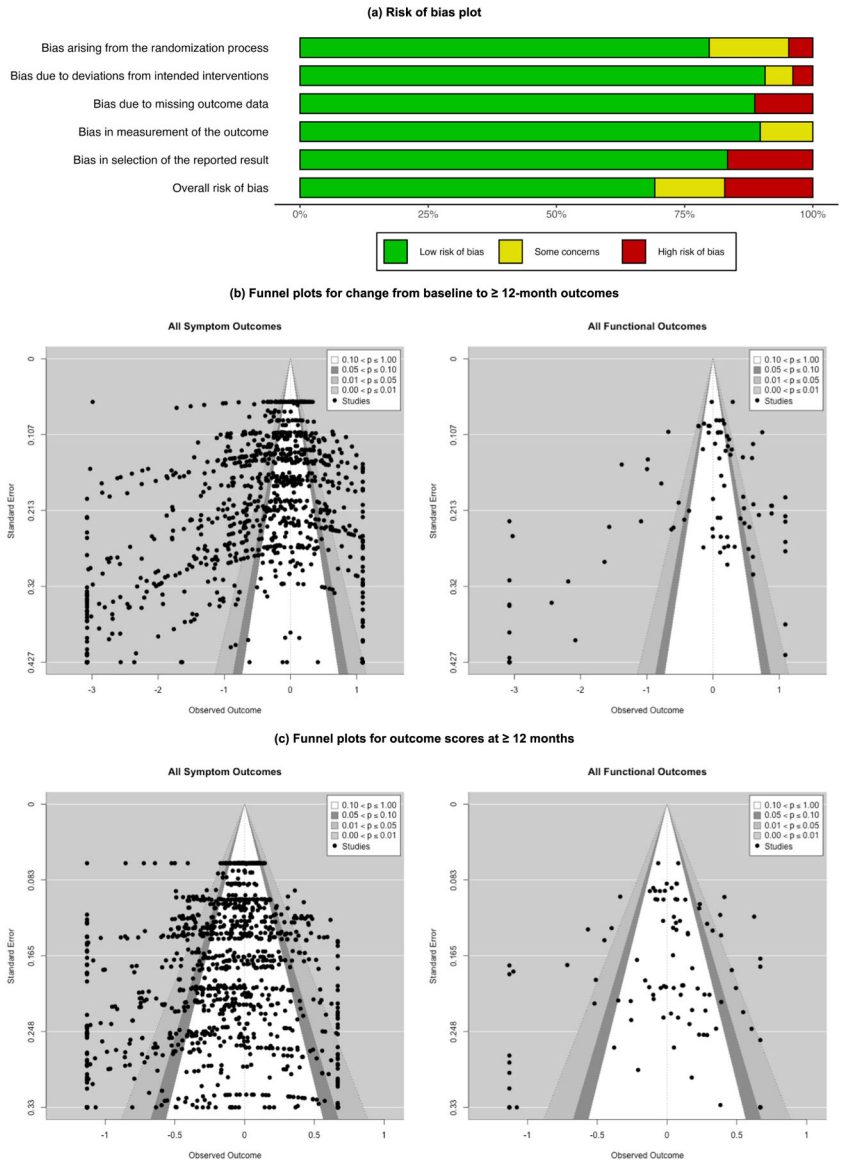


**Fig. 4.**
Risk of bias plot.

*Clin Psychol Rev*. Author manuscript; available in PMC 2025 February 24.

**Table 1**

Guided i-CBT vs. all controls on long-term follow-up outcomes.

| | Model | | Effect size | | 95 % CI | | Heterogeneity | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *k* | $k_{ES}$ | *g* | *p* | LCI | UCI | *Q* | $Q_{df}$ | $Q_p$ | $I^2$ |
| a. Predicting change in scores from baseline to 12-month follow-up | | | | | | | | | | |
| Depression | 81 | 119 | −0.31 | **.000** | −0.44 | −0.18 | 922.0 | 118 | **.000** | 88.6 |
| Anxiety | 44 | 102 | −0.44 | **.000** | −0.58 | −0.31 | 335.8 | 101 | **.000** | 74.1 |
| QOL | 56 | 103 | 0.07 | .443 | −0.10 | 0.24 | 1070.4 | 102 | **.000** | 92.2 |
| Role impairment | 28 | 54 | −0.09 | .401 | −0.29 | 0.12 | 738.5 | 53 | **.000** | 93.1 |
| Insomnia | 9 | 12 | −0.31 | .261 | −0.85 | 0.23 | 38.2 | 11 | **.000** | 77.4 |
| RNT | 10 | 19 | −0.69 | **.000** | −1.05 | −0.33 | 100.9 | 18 | **.000** | 83.1 |
| Global symptoms | 10 | 12 | −1.56 | .156 | −3.72 | 0.60 | 540.9 | 11 | **.000** | 99.5 |
| Self-efficacy | 12 | 17 | −0.10 | .672 | −0.57 | 0.37 | 138.5 | 16 | **.000** | 91.6 |
| Emotion regulation | 12 | 35 | −0.28 | .514 | −1.12 | 0.56 | 739.8 | 34 | **.000** | 98.2 |
| PTSD | 5 | 11 | −1.86 | **.000** | −2.37 | −1.36 | 180.0 | 10 | **.000** | 92.3 |
| Pain | 8 | 13 | −0.58 | .230 | −1.51 | 0.36 | 379.6 | 12 | **.000** | 98.6 |
| b. Predicting scores aggregated across 12-, 15-, 18-, 24-, 30-, 36-, 42-, 48-, and 120-month follow-ups | | | | | | | | | | |
| Depression | 113 | 194 | −0.18 | **.000** | −0.25 | −0.12 | 1108.8 | 193 | **.000** | 83.1 |
| Anxiety | 78 | 213 | −0.23 | **.000** | −0.30 | −0.15 | 1460.1 | 212 | **.000** | 84.0 |
| QOL | 74 | 163 | 0.11 | **.000** | 0.06 | 0.17 | 600.7 | 162 | **.000** | 70.5 |
| Distress | 31 | 49 | −0.12 | **.006** | −0.21 | −0.03 | 169.0 | 48 | **.000** | 74.5 |
| Role impairment | 41 | 87 | −0.18 | **.001** | −0.29 | −0.08 | 378.7 | 86 | **.000** | 84.7 |
| Insomnia | 16 | 29 | −0.31 | **.000** | −0.46 | −0.16 | 108.5 | 28 | **.000** | 67.3 |
| RNT | 14 | 32 | −0.17 | .061 | −0.36 | 0.01 | 286.7 | 31 | **.000** | 85.4 |
| Global symptoms | 12 | 14 | −0.17 | .294 | −0.48 | 0.14 | 102.4 | 13 | **.000** | 91.0 |
| Self-efficacy | 12 | 19 | −0.01 | .924 | −0.14 | 0.12 | 66.9 | 18 | **.000** | 75.3 |
| Emotion regulation | 12 | 55 | 0.26 | **.000** | 0.14 | 0.38 | 158.9 | 54 | **.000** | 60.7 |
| PTSD | 8 | 17 | −0.28 | **.000** | −0.42 | −0.14 | 30.9 | 16 | **.014** | 36.8 |
| Pain | 8 | 13 | −0.11 | .289 | −0.30 | 0.09 | 60.9 | 12 | **.000** | 80.0 |

i-CBT, internet-delivered cognitive-behavioral therapy; CI, confidence interval; *k*, number of studies; $k_{ES}$, number of effect sizes; *g*, Hedge's *g* effect size estimate; LCI, lower bound of the 95% CI; UCI, upper bound of the 95% CI; *Q*, test for heterogeneity statistic; *df*, degrees of freedom; QOL, quality of life; RNT, repetitive negative thinking; PTSD, post-traumatic stress disorder; OCD, obsessive-compulsive disorder. The outcome of the set of analyses in Part b was derived by pooling the means and *SDs* of the scores reported at follow-ups of ≥12 months after randomization.

**Table 2**

Self-guided i-CBT vs. all controls on long-term follow-up outcomes.

| | Model | | Effect size | | 95 % CI | | Heterogeneity | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***k*** | $k_{ES}$ | ***g*** | ***p*** | **LCI** | **UCI** | ***Q*** | $Q_{df}$ | $Q_p$ | $I^2$ |
| a. Predicting change in scores from baseline to 12-month follow-up | | | | | | | | | | |
| Depression | 23 | 45 | −0.51 | **.007** | −0.88 | −0.14 | 853.3 | 44 | **.000** | 99.0 |
| Anxiety | 15 | 42 | −0.24 | .055 | −0.48 | 0.01 | 553.4 | 41 | **.000** | 96.3 |
| Role impairment | 5 | 25 | −1.83 | .214 | −4.71 | 1.06 | 405.1 | 24 | **.000** | 91.0 |
| b. Predicting scores aggregated across 12-, 15-, 18-, 24-, 30-, 36-, 42-, 48-, and 120-month follow-ups | | | | | | | | | | |
| Depression | 14 | 33 | −0.09 | **.001** | −0.14 | −0.04 | 27.0 | 32 | .718 | 7.7 |
| Anxiety | 12 | 22 | −0.16 | **.000** | −0.23 | −0.09 | 11.2 | 21 | .958 | 13.8 |
| QOL | 11 | 27 | 0.02 | .834 | −0.13 | 0.16 | 61.2 | 26 | **.000** | 77.7 |
| Role impairment | 8 | 23 | −0.09 | .378 | −0.29 | 0.11 | 142.8 | 22 | **.000** | 84.3 |

i-CBT, internet-delivered cognitive-behavioral therapy; CI, confidence interval; *k*, number of studies; $k_{ES}$, number of effect sizes; *g*, Hedge's *g* effect size estimate; LCI, lower bound of the 95% CI; UCI, upper bound of the 95% CI; *Q*, test for heterogeneity statistic; *df*, degrees of freedom; QOL, quality of life. The outcome of the set of analyses in Part b was derived by pooling the means and *SDs* of the scores reported at follow-ups of ≥ 12 months after randomization.