Effects of Internet-based Cognitive Behavioral Therapy in Routine Care for Adults in Treatment for Depression and Anxiety: A Systematic Review and Meta-analysis

Anne Etzelmueller, Christiaan Vis, Eirini Karyotaki, Harald Baumeister, Nickolai Titov, Matthias Berking, Pim Cuijpers, Heleen Riper, David Daniel Ebert

Submitted to: Journal of Medical Internet Research
on: February 03, 2020

Table of Contents

Effects of Internet-based Cognitive Behavioral Therapy in Routine Care for Adults in Treatment for Depression and Anxiety: A Systematic Review and Meta-analysis

Anne Etzelmueller; Christiaan VisMSc, ; Eirini KaryotakiPhD, ; Harald BaumeisterPhD, ; Nickolai TitovPhD, ; Matthias BerkingPhD, ; Pim CuijpersPhD, ; Heleen RiperPhD, ; David Daniel EbertPhD,

Corresponding Author:
Anne Etzelmueller
Email: anne.etzelmueller@fau.de

Abstract

Background: While there is evidence for the efficacy of internet-based Cognitive Behavioral Therapy (iCBT), the generalizability of results to routine care is limited.

Objective: This study systematically reviews effectiveness studies of guided iCBT interventions for the treatment of depression and/or anxiety.

Methods: The acceptability (uptake, participant´s characteristics, adherence, and satisfaction), effec-tiveness and negative effects (deterioration) of non-randomized pre-post designs conducted under routine care conditions were synthesized using systematic review and meta-analytic approaches.

Results: 19 studies including 30 groups were included in the analysis. Despite high heterogeneity, individual effect-sizes of investigated studies indicate clinically relevant changes, with effect sizes ranging from Hedges' g = 0.42 - 1.88 with a pooled effect of 1.78 for depression and 0.94 for the anxiety studies. Uptake, participant characteristics, adherence and satisfaction indicate a moderate to high acceptability of the interventions. The average deterioration across studies was 3.1%.

Conclusions: This study provides evidence supporting the acceptability and effectiveness of guided iCBT for the treatment of depression and anxiety in routine care. Given the high heterogeneity between interventions and contexts, health care providers should select interventions that have been proven in randomized controlled clinical trials. The successful application of iCBT may be an effective way of increasing health care in multiple contexts. Clinical Trial: This meta-analysis was registered at PROSPERO under the registration number CRD42018095704.

(JMIR Preprints 03/02/2020:18100)
DOI: https://doi.org/10.2196/preprints.18100

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
✓ **No, I do not wish to publish my submitted manuscript as a preprint.**
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Effects of Internet-based Cognitive Behavioral Therapy in Routine Care for Adults in Treatment for Depression and Anxiety: A Systematic Review and Meta-analysis

Anne Etzelmüller [a, b, c, d]

Christiaan Vis [c, d]

Eirini Karyotaki [c, d, e]

Harald Baumeister [f]

Nickolai Titov [g]

Matthias Berking [a]

Pim Cuijpers [c,d]

Heleen Riper [c, d, h]

David Daniel Ebert [a, b, c, d]

[a] Clinical Psychology and Psychotherapy, Friedrich-Alexander University Erlangen Nueremberg

[b] GET.ON Institute GmbH, Germany

[c] Department of Clinical, Neuro-, & Developmental Psychology, Faculty of Behavioural and Movement Sciences, VU Amsterdam, Amsterdam, Netherlands

[d] Mental Health, Amsterdam Public Health Research Institute, Amsterdam, Netherlands

[e] Department of Global Health and Social Medicine, Harvard Medical School, USA

[f] Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University

[g] eCentre Clinic Department of Psychology, Macquarie University, Sydney, Australia

[h] Community Mental Health Centre GGZ inGeest, Amsterdam, the Netherlands

Corresponding author: Anne Etzelmüller, Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander University Erlangen-Nuremberg, Naegelbachstrasse 25a, 90453 Erlangen, Germany. Email: anne.etzelmueller@fau.de, Tel.: 0049 9131 85 67575.

Abstract

Background. While there is evidence for the efficacy of internet-based Cognitive Behavioral Therapy (iCBT), the generalizability of results to routine care is limited. **Objective.** This study systematically reviews effectiveness studies of guided iCBT interventions for the treatment of depression and/or anxiety. **Method.** The acceptability (uptake, participant´s characteristics, adherence, and satisfaction), effectiveness and negative effects (deterioration) of non-randomized pre-post designs conducted under routine care conditions were synthesized using systematic review and meta-analytic approaches. **Results.** 19 studies including 30 groups were included in the analysis. Despite high heterogeneity, individual effect-sizes of investigated studies indicate clinically relevant changes, with effect sizes ranging from Hedges' g = 0.42 - 1.88 with a pooled effect of 1.78 for depression and 0.94 for the anxiety studies. Uptake, participant characteristics, adherence and satisfaction indicate a moderate to high acceptability of the interventions. The average deterioration across studies was 2.9%. **Conclusion.** This study provides evidence supporting the acceptability and effectiveness of guided iCBT for the treatment of depression and anxiety in routine care. Given the high heterogeneity between interventions and contexts, health care providers should select interventions that have been proven in randomized controlled clinical trials. The successful application of iCBT may be an effective way of increasing health care in multiple contexts.

Keywords: Internet-based interventions; depression; anxiety; effectiveness; acceptability; routine care.

Trial registration: This meta-analysis was registered at PROSPERO under the registration number CRD42018095704.

Introduction

Depressive and anxiety disorders are common mental health problems associated with significant suffering, impairment and reduction in quality of life [1,2]. Both disorders lead to considerable socioeconomic costs through decreased work productivity and higher utilization of healthcare services [3,4].

Despite the proven effectiveness of psychotherapy in the treatment of depression and anxiety [5], the provision of evidence-based treatments depicts a constant challenge given barriers such as the shortage, uneven distribution of trained providers, delayed treatment provision, and inadequacy of treatment [6,7]. Furthermore, research on patient's preferences has shown that many do neither make use of psychotherapeutic treatments nor do they receive psychopharmacological treatment [7]. Using the internet to provide psychotherapeutic interventions may increase the coverage of usual care services [8,9] by providing highly accessible and scalable interventions reaching people who cannot be reached otherwise. Recent research suggests that internet-based CBT (iCBT) with therapeutic guidance is effective for the prevention [10,11] and treatment [12–15] of common mental disorders. Systematic reviews on studies were also able to show comparable effects to face-to-face treatments in adults [16,17]. In a recent meta-analysis, Romijn and colleagues [13] showed that iCBT interventions for anxiety disorders can also have significant effects obtained in trials implemented in clinical care. They also found that effects were smaller in samples recruited in clinical practice than within samples recruited with an open recruitment method compared to waitlist-control groups [13], which raises the question of the effects of iCBT when implemented in routine practice.

While randomised controlled trials (RCTs) are considered the gold-standard in exploring the efficacy of mental health interventions, the idealized and controlled nature of these trials limits the generalisability of findings to routine care populations [18]. RCTs maximize the internal validity, to make sure that the effect found can be attributed to the investigated intervention [19,20]. Thus, RCT findings are restricted by controlled protocols, explicit eligibility criteria, and patient recruitment and randomization procedures. RCTs provide a highly structured environment, which is considered to possibly have an adherence-fostering effect [21,22]. The efficacy derived from randomised-controlled trials of internet-based interventions might be overestimated for what can be expected when implementing in routine care limiting the knowledge base for routine clinical practice [20].

Hence, after establishing the efficacy of an intervention and its subsequent implementation, so-called “Phase IV clinical trials” should follow investigating benefits when implemented as well as potential negative effects implemented [23,24]. Thus, the investigation of the effectiveness of iCBT under routine care conditions depicts an important part of the evaluation of these services before wide-scale adoption.

Andersson and Hedman [25] reported on the effectiveness of iCBT within four controlled trials and eight open studies for a multitude of mental health problems indicating that it might be possible to replicate the findings of controlled efficacy trials on guided ICBT in clinical practice. However, in that review, both routine care and randomized controlled trials were included and only eight studies reported effects when the service was delivered under routine conditions. Recently, Andrews and colleagues [15] reported on results of computer-based treatments of depression, panic disorder, generalized anxiety disorder and social phobia in randomized trials. They also identified eight studies on internet-based treatments in routine clinical practice when delivered outside of a randomized trial reporting an average effect size of g = 1.07 across all four disorders [15]. However, since then, many more studies have been published. Also, this review did not specifically try to identify non-randomised trials, possibly leading to unidentified articles. Additionally, they did not provide disorder-specific results, specific results on guided treatments by mixing guided and unguided treatments and did not investigate the acceptability and potential negative effects.

The aim of this study was to examine the effects of guided iCBT for the treatment of depression and anxiety under routine care conditions on symptom change, acceptability (uptake, participant´s characteristics, adherence, and satisfaction), average and predictors of negative effects (deterioration, side-effects).

Methods

We report this meta-analysis in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Appendix A). This meta-analysis was registered at PROSPERO (registration number: CRD42018095704).

We searched PubMed, PsychINFO and the Cochrane library. We used index terms and text words associated with depression and anxiety, internet interventions and routine care (for a full search string, the reader is referred to Appendix B). Furthermore, we contacted experts in the field

to ask whether they were aware of studies that we did not identify through our systematic literature searches. Further, we conducted reference tracking on the identified studies and previous meta-analyses on the field [5,14,15,26]. The resulting hits of our literature searches were screened on titles and abstracts by two independent reviewers (AE and CV). Studies considered as potentially relevant were screened on full text by the same reviewers independently. In case of disagreement, the opinion of a third senior reviewer (DDE) was sought.

Inclusion criteria

We included studies, which (a) examined the effectiveness of a guided or blended internet-based Cognitive Behavioral Therapy in (b) treating adults with depressive and/or anxiety symptoms (c) under routine care conditions (d) in a pre-post design. We followed the inclusion of adults and older adolescents (> 16 of age) within the treatment provision for adults as reported in the original studies.

We defined routine care studies as effectiveness studies, which are conducted as non-randomized clinical trials in settings being equal to or representative of routine practice [27]. The definition of routine care differs between countries and health care systems and describes the established way of working at the time of the original study. Depression and anxiety symptoms had to be established based on cut-off scores on self-report outcome measures, clinical diagnosis, or expert opinion. The definition of anxiety symptoms is based on the DSM-IV classification criteria for anxiety disorders. Further, the interventions were considered as guided when the guidance was related to the therapeutic content [28] and as blended when the internet-based intervention was combined with face-to-face elements in one integrated standardized treatment protocol [26]. Guidance could be delivered via email, a secure message system, telephone, or face-to-face contact and via video or face-to-face contact in blended treatments. Finally, both disorder-specific and transdiagnostic interventions (targeted at both depression and anxiety simultaneously) were included.

Exclusion criteria

We excluded studies, which did not (a) focus primarily on anxiety or depression or (b) provide sufficient data for the calculation of the effect sizes. Studies were also excluded if (c) the service had only been provided as part of a research study, (d) the study could be considered as a feasibility or pilot trial, or (e) patients were randomized at an individual level. However, cluster randomized trials were considered eligible, in which randomization took not place on an individual level, but for

example on a health care institution level. For the definition of feasibility and pilot trials, we followed the NETSCC definition of pilot and feasibility trials [29]as recommended by Arain, Campbell, Cooper, & Lancaster (2010). Feasibility trials were defined as “pieces of research done before a main study” (designed around the research question “Can this be done?”) and pilot studies are defined as a version of the main study that is “run in miniature to test whether the components of the main study can all work together” [30]. Also, we only included studies published in English, German or Dutch language.

Data extraction.

We extracted data related to study and iCBT service-related characteristics, acceptability, effects on symptom change, negative effects, and data related to the risk of bias of reported results.

Study characteristics included the year of publication, the country in which the study was conducted, the year of data collection, sample size and eligibility criteria (establishment of depression/anxiety diagnosis at baseline [standardized clinical Interview, cut-off on standardized questionnaire, clinical judgement], inclusion of severe cases [yes/no], exclusion of cases with suicidal ideation [yes/no]) and approach to data analysis [ITT/completer].

ICBT service-related characteristics included intervention name, the symptoms targeted (depression and/or anxiety), if it was a blended treatment (yes/no), evidence base for the used intervention, (positive results based on at least one randomized clinical trial [yes/no]) and whether it was a symptom-specific or transdiagnostic treatment. We also included the recruitment pathway (open community, clinical referral, both), the number of planned intervention modules, guidance focus (content-focused, motivational-focussed and administrative-focussed), guidance delivery format (synchronous vs. asynchronous, within the treatment platform vs. outside, e.g. by email) and guidance moment (as a reaction to an action of the participant [e.g. after the participant finished a session, as a reaction to a non-response], or planned in different intervals [e.g. weekly or bi-weekly]). Furthermore, we included guide’s professional training (psychotherapist, psychiatrist, GP, psychologist, psychological registrar, nurse, coach [with lived experience]), training of professionals in iCBT (yes/no), supervision of professionals by a trained clinician (yes/no), as well as the planned and actual intensity of guidance in minutes and if there was a guidance manual provided (yes/no). Additional information on whether a standardized procedure in case of symptom deterioration and

crisis (yes/no) has been established was included.

Acceptability data were extracted with regard to uptake (the number of people screened for the service, people included and participants starting the treatment), patient characteristics (age, gender), average symptom severity at baseline, adherence (i.e. number of completed modules), mean treatment duration in weeks and participant´s satisfaction. Negative effects were extracted with regard to average effects on symptom deterioration, other side effects, and report of specific subgroups at risk for symptom deterioration.

Two reviewers (AE and CV) extracted the data independently and data sets were merged. Differences and points of uncertainty were discussed and checked by returning to the original article and in some cases to the authors of the respective article.

Risk of Bias Assessment.

Assessing the quality of naturalistic observational studies is challenging as there is no widely accepted tool in doing so [31]. Moreover, established guidelines of the quality assessment of non-randomised trials are only partially applicable as they assume comparisons of interventions (ROBINS-I [32]). Thus, in this study, we selected and adapted criteria from two quality assessment tools [32,33] and adapted them to this study's purposes to evaluate the Risk of Bias of the included studies. For the present Risk of Bias assessment, we discussed the aforementioned assessment tools among all co-authors of this manuscript and derived the analysis criteria described in Appendix C. As a result, we evaluated a) researcher allegiance (defined as the first or last author of the study also being the first or last author of the intervention development or efficacy paper), b) confounding introduced by patient's participation in other treatments, c) confounding introduced by significant confounding variables identified within the individual study (meaning any predictors included such as age, guidance or recruitment pathway), d) selection bias introduced by the study population (i.e. have the studies only reported on completer data) and e) selective outcome reporting in comparison to the study protocol or diagnostic measures administered as mention in the original studies methods section. A description of the Risk of Bias assessment and its operationalisation can be found in Appendix C. With regards to "Researcher Allegiance" we chose the above definition after consideration among the authors, and evaluated a study as at high risk of re- searcher allegiance when the first or last author of the study were also involved in the development of the treatment manual of the psychotherapy involved or the reporting on the interventions efficacy.

While the validity of other indicators has been questioned, the involvement of a researcher in developing the treatment under investigation can be considered a valid indicator of potential researcher allegiance [34].

Two reviewers evaluated the quality of the included studies independently (AE and CV). Any disagreement between reviewers was solved by thorough discussion. In case the disagreement could not be resolved, a third senior reviewer was consulted (DDE).

Statistical Analysis.

Our primary outcome was the reduction of depressive or anxiety symptoms from pre- to post-test assessment. We calculated the difference in depression and anxiety symptoms between pre-post assessment divided by the weighted, pooled standard deviation (Hedges' *g*). We have chosen Hedges' *g* because it allows for small sample size bias correction [35]. Because we expected considerable heterogeneity among the studies, we used the random effects model. As a rule of thumb, effect sizes of 0.8 can be viewed as large, 0.5 as moderate, and 0.2 as small [36]. In our main analysis, we included mixed depression and/or anxiety studies into the separate depression and anxiety data sets. Statistical analysis was conducted using the Comprehensive Meta-analysis program (version 2.2.2), pooled proportions were calculated with R [37] package "meta" [38].

To calculate heterogeneity, we used the I^2-statistic and its 95% confidence intervals as an indicator of heterogeneity in percentages. Heterogeneity was interpreted as low, moderate and high when 25%, 50% and 75% respectively.

We also included the correlations of the used pre-post measures using the mean of 0.76 where none was provided for depression and 0.59 for anxiety following Balk and colleagues [39]. We also conducted sensitivity analyses for correlations set to 0.00, 0.75 and 0.99 to examine the robustness of our findings [40]. We also calculated the prediction interval, which estimates where the true effects are to be expected for 95% of similar studies that might be conducted in the future [41].

Since we expected high heterogeneity, we conducted several subgroup analyses to investigate its possible sources. The examined subgroups were related to method of analysis, time to post assessment, recruitment pathways, disorders, guidance moment (specific timing or as a reaction), guidance modality (email, message, synchronous), guide profession (with or without specific CBT training), supervision provided [yes/no], guide training provided [yes/no], intervention manual provided [yes/no], approach to data analysis [ITT/completer] and diagnostic method [interview/questionnaire]. Subgroup analyses were only carried out with regard to effects on

symptom change. We utilized the mixed effects model, testing pooled studies within subgroups with random effects models while testing for significant differences between those subgroups with fixed effects models. We only conducted subgroup analysis if the number of studies per category was not smaller than three. If necessary, we combined pre-defined subgroups to achieve the necessary group size.

Finally, we conducted meta-regression analyses for the continuous variables, examining the duration of the treatment as a predictor of treatment outcome as well as guidance time, number of contacts, number of sessions completed, and the percentage of treatment completers

Regarding uptake, we calculated the proportion of a) included people based on the number of people screened, b) starters based on the number of people being screened and c) starters based on the number of people included. Adherence was analysed by calculating the percentage of modules completed based on the average number of sessions that were completed by the participants divided by the planned total number of sessions. We also coded the percentage of intervention completers for a 100% completion rate. Additionally, we pooled the age and gender distribution as well as participant satisfaction extracted from original studies. Furthermore, we pooled the percentage of individuals reported to showing symptom deterioration (defined as negative reliable change on the reported outcome), the deterioration rates, reported in the original study.

Publication bias was examined by inspecting the funnel plot [42] and conducting the Egger's test of the intercept with a one-tailed significance level a = .05 [43]. Also, we used Duval and Tweedie's trim and fill procedure [44] to adjust the effect size for missing studies.

Results

Study selection.

The systematic literature search was performed on January 30th, 2019. This search resulted in 25 447 citations. After removal of duplicates, 19 316 citations remained for the title and abstract

assessment and 174 after the exclusion due to title and abstract. Seventeen studies fulfilled the eligibility criteria. The full references of included studies are listed in Appendix D. The study selection process is described in Figure 1.

Study characteristics.

Table 1 presents the characteristics of included studies. Seventeen studies (n = 12 096 participants) reporting on the outcomes in the treatment for depression and anxiety were included. Seven of the 17 studies reported multiple groups, of which five combined results on multiple treatments in the published study and two reported distinct forms of guidance within the same treatment and setting without randomising patients on an individual level. Of the resulting 30 groups, eight groups focused on depression, 17 on anxiety, and five on both depression and/ or anxiety. We included the studies reporting on both depression and/ or anxiety in both, the depression and anxiety analysis. 46.7% of the studies (k = 14/30, k_{Dep} = 4, k_{Anx} =10) administered diagnostic interviews, 36.7% self-reports (k = 11/30; k_{Dep} = 6/13, k_{Anx} = 7/22), and 16.7% (k = 5/30; k_{Dep} = 3/13, k_{Anx} = 2/22) clinical judgement in their diagnostic process. 30.0% percent of studies (k = 9/30; k_{Dep} = 4/13, k_{Anx} = 5/22) administered a cut-off criterion for the self-report questionnaires. Ten studies (33.3%, k = 10/30; k_{Dep} = 6/13, k_{Anx} = 5/22) included patients who did not meet criteria for a clinical diagnosis of depression or anxiety. 9.1% anxiety studies (k = 2/22) excluded cases with severe symptom severity, all depression studies allowed patients with severe depression severity to be included. 40.0% percent of studies (k = 12/30; k_{Dep} = 4/13, k_{Anx} = 9/22) specifically stated that the patients had to be diagnosed with a clinical depression and/ or anxiety disorder to follow the iCBT intervention. The rest of the studies did not specify whether the patients had clinical depression and/ or anxiety. Seventy-three percent of the studies (k = 22/30; k_{Anx} = 9/13, k_{Dep} = 18/22) stated that suicidal ideation or intent was a reason for excluding the patient from the service.

Figure 1 *Study inclusion*

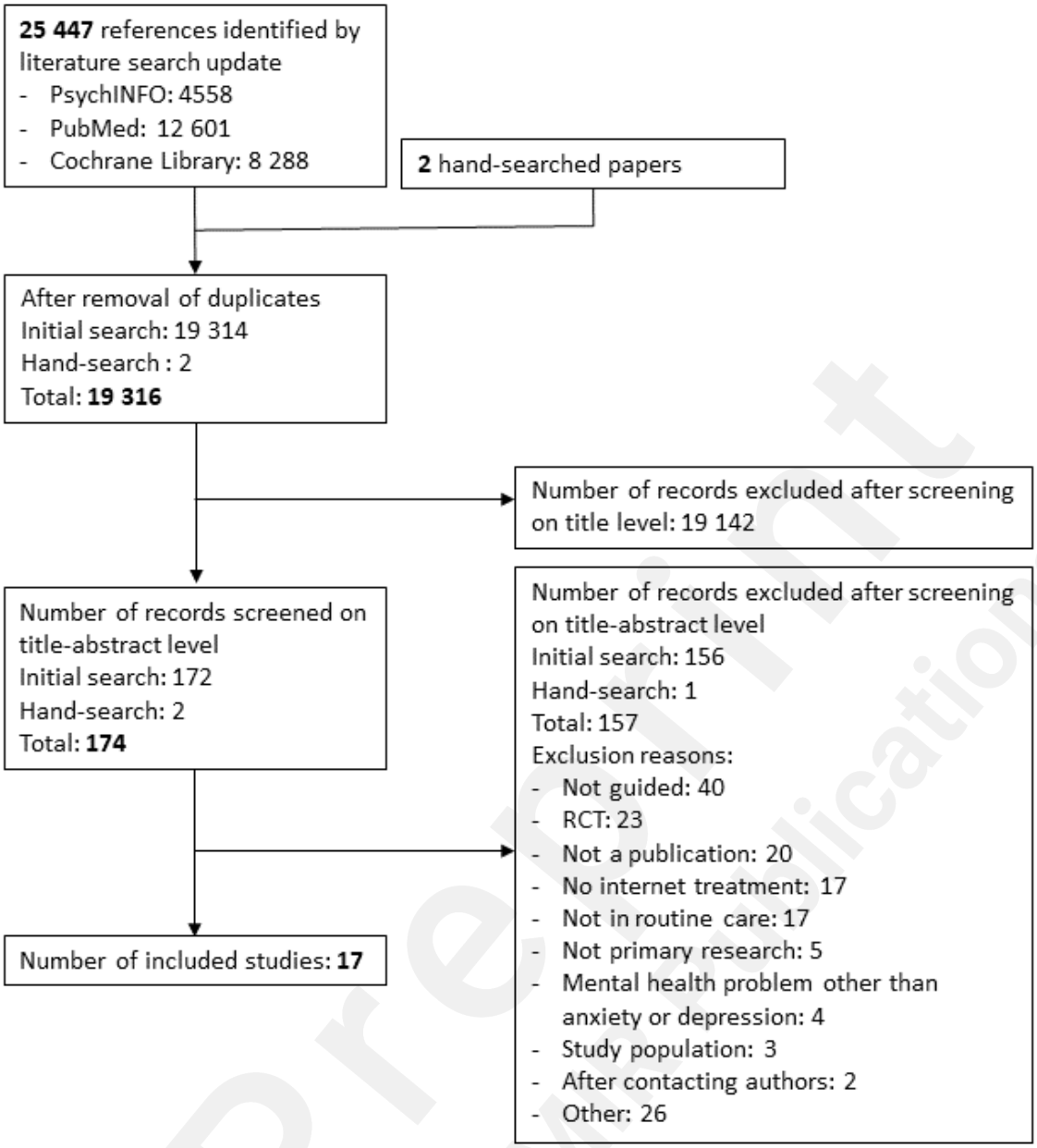


Table 1 *Study characteristics*

Publication	Sub-Study	Year of publication	Data collection [a]	Country	Sample size	Diagnosis conducted	criterionDiagnostic	e casesInclusion of	[b]Exclusion: suicidal
Aydos et al., (2009)		2009	NA	AUS	17	Interview (MINI)	Clinical	Yes	No
Alaoui et al., (2015)		2015	2009 - 2013	SE	653	Interview (MINI)	Clinical	Yes	Yes
Etzelmueller et al., (in prep.)		NA	2014 - 2017	DE	349	Self-report (PHQ8 > 10)	Clinical and Sub-clinical	No	No
Gellatly et al., (2018)		2018	2013 - 2015	UK	724	Clinical judgement	Caseness [c]	No	No
Hadjistavropoulos et al. (2014)	GAD [d]	2014	2010 - 2013	CA	107	Interview (MINI) + GAD7>5	Clinical and Sub-clinical	No	Yes
	Depression				80	Interview (MINI) + PHQ>5			
	Panic Disorder				25	Interview (MINI) + PDSS-SR > 8			
Hadjistavropoulos et al. (2016)	Specialised Care	2016	2013 - 2015	CA	260	Self-report (K10 ≥ 17)	Clinical	No	Yes
	Non-Specialised Care				198				
Hedman et al. (2013)		2013	2007 - 2012	SE	1203	Interview (MINI)	Clinical	No	No
Hedman et al. (2014)		2014	2007 - 2013	SE	570	Interview (MINI)	Clinical	No	No
Marks et al. (2003)	Phobia/ Panic	2003	NA	UK	27	Clinical judgement (ICD10)	Clinical	No	Yes
	Depression [e]				38				
	Anxiety/ depression				33				
	OCD [f]				9				
Mathiasen et al. (2018)	Depression	2018	2016 - 2017	DK	60	Interview	Clinical	No	Yes
	Anxiety				143				

Publication	Sub-Study	Year of publication	Data collection [a]	Country	Sample size	Diagnosis conducted	criterionDiagnostic	e casesInclusion of	[b]Exclusion: suicidal
Morrison et al. (2014)		2014	2012	UK	12	Self-report and clin. judgem. [g]	Caseness [c]	No	No
Nordgreen et al. (2018)		2018	2014 - 2016	NO	124	Interview (MINI)	Clinical	No	Yes
Nordgreen et al. (2018b)		2018	NA	NO	169	Interview (MINI)	Clinical	No	Yes
Ruwaard et al. (2012)	Depression	2012	2002 - 2008	NL	405	Interview (NA)	Clinical	No	Yes
	Panic Disorder				136				
	PTSD [h]				477				
Shandley et al. (2008)	GP-guided	2008	NA	AUS	51	Self-report and Interview	Clinical	No	No
	Therapist-guided				41				
(Titov et al., 2017)	Depression	2017	2013 - 2016	AUS	5427	Self-report	Principal complaint	No	Yes
	Depression [i]				516				
	OCD [f]				69				
	PTSD [h]				137				
(Yu et al., 2018)		2018	NA	US	63	Self-report (GAD7 ≥ 5)	Clinical	No	Yes

Note. Full references are available in Appendix D.

[a] Data collection period

[b] Exclusion of cases with suicidal ideation

[c] Caseness for PHQ-9 refers to a person reporting scores of 10 on the PHQ-9

[d] Generalised Anxiety Disorder

[e] Transdiagnostic treatment for depressed

[f] Obsessive compulsive disorder

[g] Participants were initially identified as suitable to receive a low-intensity intervention for depression or low mood through triage of a patient's self-assessment form by team leaders, all of whom were qualified CBT therapists. Patients then had an initial assessment with a [Psychological Wellbeing Practitioner] who considered a person's suitability for MindBalance in reference to the patient's identified difficulties, goals and the studies inclusion and exclusion criteria (Inclusion: To receive treatment of depression with little or no comorbid anxiety, appropriate for guided self-help

in a primary-care setting as determined by current [...] procedures)
[h] Post-traumatic stress disorder
[i] Depression treatment for older adults

iCBT service-related characteristics.

26.3% of studies (*k* = 5/19) used transdiagnostic interventions, all others utilized disorder specific interventions. We did not identify any blended treatments.

31.6% of the included studies (*k* = 6/19) involved clinical referrals in their service pathway, 26.3% (*k* = 5/19) did not involve referrals but solely included patients through the general community, whereas 42.1%, (*k* = 8/19) recruited in both, a community and clinical setting.

On average, iCBT treatments included 8.00 sessions (SD = 2.62, *k* = 26; depression: *k* = 11, *M* = 8.09, *SD* = 2.84; anxiety: *k* = 19, *M* = 8.00, *SD* = 2.81).

With regard to guidance, 46.7% of the studies (*k* = 14/30; k_{Dep} = 5/13, k_{Anx} = 11/22) stated that guidance focused mainly on motivational and 16.7% (*k* = 5/30; k_{Dep} = 4/13, k_{Anx} = 4/22) on administrative aspects. All included studies provided feedback on the content of participants completed sessions. 73.3% of studies (*k* = 22/30; k_{Dep} = 10/13, k_{Anx} = 14/22) used asynchronous contact methods for the communication between participant and guide, 30.0% of studies (*k* = 9/30; k_{Dep} = 7/13, k_{Anx} = 4/22) used build-in message systems and 16.7% (*k* = 5/30; k_{Dep} = 1/13, k_{Anx} = 5/22) emails. 23.3% of studies (*k* = 7/30; k_{Dep} = 3/13, k_{Anx} = 7/22) used synchronous contact via telephone contacts, of which one would also use face-to-face contacts. 16.7% of studies (*k* = 5/30; k_{Dep} = 1/13, k_{Anx} = 4/22) stated that they provided feedback as a reaction following a participant's action and 30.0% (*k* = 9/30; k_{Dep} = 4/13, k_{Anx} = 6/22) in specific time interval, weekly or bi-weekly.

23.3% of studies (*k* = 7/30; k_{Dep} = 4/13, k_{Anx} = 6/22) only involved guides not trained in CBT, while the other studies including specifically trained professionals, such as psychotherapist, psychiatrist, GPs, or psychologist. 40.0% of studies (*k* = 12/30; k_{Dep} = 7/13, k_{Anx} = 10/22) stated that they provided specific training for the provision of the iCBT intervention to the guides, and 63.3% (*k* = 19/30; k_{Dep} = 9/13, k_{Anx} = 13/22) provided supervision to the those guiding participants. 26.7% studies (*k* = 8/30; k_{Dep} = 4/13, k_{Anx} = 4/22) reported having provided an iCBT intervention manual. The average reported guidance time was 133.49 minutes (*k* = 11, 95%-CI = 92.85 – 204.21; depression: *k* =4, *M* = 101.10 minutes, 95%-CI = 49.53 – 152.67; anxiety: *k* = 8, *M* = 142.76 minutes, 95%-CI = 109.80 – 175.72). The pooled results are depicted in Table 2.

Nine studies (47.4%, *k* = 19) reported safety measures in case of suicidality, suicidal

ideation or severe symptom deterioration. They mention monitoring systems (*k* = 2), risk alerts (*k* = 1) or reviewing the participants messages (*k* = 2) as ways to identify risk. Procedures were triggered in case of suicidal ideation or suicidality (*k* = 9), inactivity or lack of progress (*k* = 2), or an increase in symptoms (*k* = 2). Ways to mitigate the risk included contacting the participant via telephone (often in contrast to the usual messaging, *k* = 4), structured risk assessments (*k* = 1), referred to another service (v = 6), and the development of a safety plan together with the participant (*k* = 1). Information is depicted in Appendix E and F on the iCBT service-related characteristics.

Table 2 Pooled results of iCBT service- and acceptability-related outcomes guidance time, age, gender, completed sessions, completed components and deterioration rates

		N [a] studies	Pooled mean / %	95% CI [b]	Range
Guidance Time (in minutes)	All studies	14	254.0	177.0 - 331.2	43.0 - 378.6
	Depression studies	5	322.3	67.7 - 576.9	43.0 - 183.0
	Anxiety studies	10	197.9	130.9 - 264.8	43.0 - 378.6
Age	All studies	29	38.3	37.2 - 39.4	29.8 - 43.5
	Depression studies	12	39.00	37.8 - 40.2	29.0 - 41.7
	Anxiety studies	21	37.8	36.5 - 39.2	29.8 - 43.5
Gender (female, in %)	All studies	23	65.4	57.2 - 72.8	22.2 - 91.7
	Depression studies	11	70.1	55.7 - 81.4	22.2 - 91.7
	Anxiety studies	17	64.3	56.1 - 71.6	22.2 - 86.0
Average percentage of sessions completed	All studies	14	60.6	57.2 - 72.8	16.7 - 90.0
	Depression studies	5	62.6	61.2 - 63.9	16.7 - 90.0
	Anxiety studies	10	57.3	56.1 - 58.4	16.7 - 74.3
Average percentage of participant completing all treatment components	All studies	26	61.0	55.3 - 66.9	27.3 - 82.6
	Depression studies	12	62.8	55.1 - 70.0	44.0 - 82.6
	Anxiety studies	18	61.7	53.5 - 69.3	27.3 - 82.0
Deterioration	All studies	14	2.9	1.9 - 4.3	1.0 - 16.6
	Depression studies	5	2.5	2.2.-2.9	1.0 - 12.5
	Anxiety studies	9	3.1	1.6 - 5.9	1.0 - 16.6

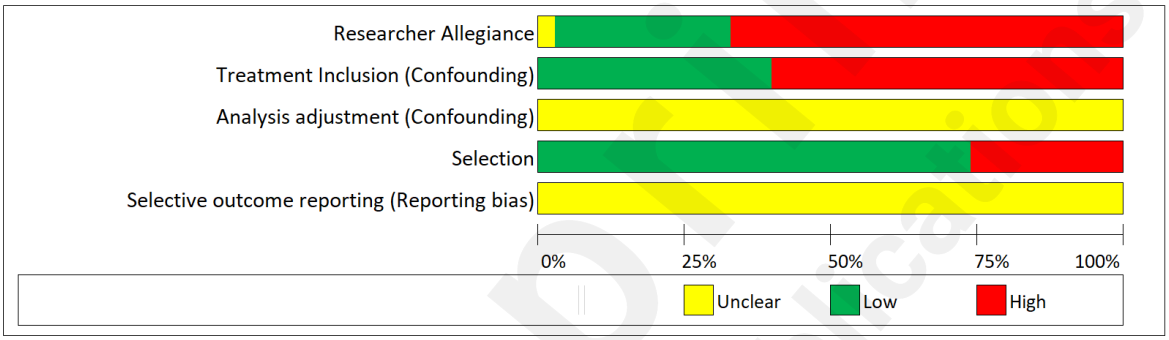
[a] Number of studies
[b] Confidence Interval

an increase in symptoms (*k* = 2). Ways to mitigate the risk included contacting the participant via telephone (often in contrast to the usual messaging, *k* = 4), structured risk assessments (*k* = 1), referred to another service (v = 6), and the development of a safety plan together with the participant (*k* = 1). Information is depicted in Appendix E and F on the iCBT service-related characteristics.

Risk of bias of included studies.

The quality of the included studies varied. Sixty-seven percent of the studies (*k* = 20/30) were rated with a high risk of bias on *Researcher Allegiance*. Sixty-three percent of the studies (*k* = 19/30) did not exclude patients that were participating in other psycho-therapeutic treatments (*Treatment Inclusion Confounding)* and none of the studies reported on the adjustment for confounders in the data analysis. Intention-to-treat data could be extracted from 73.3% of the studies (*k* = 22/30) and none of the studies were preceded by a published study protocol. The risk of bias assessment is depicted in Figure 2.

Figure 2 Risk of bias assessment



iCBT service acceptability.

Acceptability data on uptake, participant characteristics across studies, adherence and participant satisfaction were pooled. All acceptability results are depicted in Appendix G and H. The pooled results are depicted in Table 2.

Uptake.

The average proportion of included people based on the number of people screened was 70.2% (*k* = 6, 95%-CI = 8.4% - 98.4%, *range* = 0.6% to 76.0%), the proportion of starters based on the number of people being screened was 48.0% (*k* = 10, 95%-CI = 16.9% - 80.8%, *range* = 0.3% to 96.2%), and the proportion of starters based on the number of people included was 73.0% (*k* = 7, 95%-CI = 51.0% - 87.6%, *range* = 40.6% to 95.9%).

Participant characteristics.

The pooled percentage of female participants was 65.4% (*k* = 23, 95%-CI = 57.2%- 72.8%; depression: *k* = 11, *M* = 70.1%, 95%-CI = 55.7%-81.4%; anxiety: *k* = 17, *M* = 64.3%, 95% -CI = 56.1%-71.6%). The mean age across studies was 38.30 (*k* = 29, 95% -CI = 37.22-39.37; depression: k = 12, *M* = 38.96, 95% -CI = 37.77-40.15; anxiety: *k* = 21, *M* = 37.83, 95% -CI =

36.47-39.20).

Adherence.

The average percentage of sessions completed was 61.2% (*k* = 14, 95%-CI = 54.9% - 67.5%; depression: *k* = 5, *M* = 62.6%, 95%-CI = 61.2%-63.9%; anxiety: *k* = 10, *M* = 57.3%, 95% -CI = 56.1%-58.4%). The percentage of participant completing all treatment components was 61.3% (*k* = 26, 95%-CI = 55.3%-66.9%; depression: *k* = 12, *M* = 62.8%, 95%-CI = 55.1%-70.0%; anxiety: *k* = 18, *M* = 61.7%, 95% -CI = 53.5%-69.3%).

Participant satisfaction.

Of the seventeen studies, *k* = 10 (58.8%) studies reported the participant's satisfaction. Participant satisfaction outcomes were reported inconsistently, using varying measures and different reporting forms. Therefore, those data could not be pooled but the detailed results and the data extracted on patient satisfaction are depicted in Appendix H. Within those studies reporting participant's satisfaction, five studies reported a high and four a very high participant's satisfaction.

iCBT effects on symptom change.

Depression.

Effect sizes for changes in depression severity ranged from 0.66 to 1.88 (*Hedges' g*, *k = 13 studies*), with one study (7.7%) reporting a moderate and 12 (92.3%) a large effect size.

The average pre-post effect size of all depression treatments was *g* = 1.18 (95% CI = 1.06 – 1.29), which can be considered a large effect. Heterogeneity was significant and high (I^2 = 95%; 95% CI = 94-97; *P* < .001). The prediction interval is 0.74 to 1.62, and we can expect that in 95% of all populations, the true effect size will fall in this range.

Details of these results are shown in Figure 3 and Table 3.

In this analysis, the pre-post measurement correlation was set to the actual pre-post correlation of the measure (between 0.36 and 0.78). Sensitivity analysis, with correlations set to 0, 0.75 and 0.99, resulted in comparable effect sizes ($g_{Corr=0}$ = 1.24, $I^2_{Corr=0}$ = 86, 95% CI 78-91; *P* < .001; $g_{Corr=.75}$ = 1.16, $I^2_{Corr=.75}$= 96, 95% CI 94-97; *P* < .001), with $g_{Corr=0.99}$=0.75 ($I^2_{Corr=.99}$ = 100, 95% CI 100-100; *P* < .001) resulting in the smallest effect size.

Both the visual inspection of the funnel plot and Egger`s test (*P* = .90) did not indicate a potential publication bias.

We found five studies to be outliers, defined as not overlapping with the 95% CI of the pooled estimate. Removing these studies [45–47], and the depression group in Ruwaard and

colleagues study [48], from the analysis did not result in meaningful changes regarding effect sizes (g = 1.18, 95% CI = 1.09 – 1.26), but reduced heterogeneity (I^2 = 75%; 95% CI = 42 - 86; P < .001). Removing the mixed anxiety and depression studies did also not result in a relevant changes size (g = 1.28; 95% CI = 1.13 – 1.44; I^2=97%; 95% CI = 95-98; P < .001).

Anxiety.

For the included anxiety studies (k = 20), effect sizes ranged from 0.42 to 1.38 (*Hedges' g*), with one study (5.0%) reporting a small, six (30.0%) a moderate and 13 (65.0%) a large effect size.

The average pre-post effect size (*Hedges' g*) of all anxiety interventions, including the interventions that targeted both at anxiety and/ or depression, was g = 0.94 (95% CI = 0.83 – 1.06), which is considered a large effect. Heterogeneity was high (I^2=89, 95% CI = 84 - 92; P < .001). The prediction interval is 0.44 to 1.44, and we can expect that in 95% of all populations, the true effect size will fall in this range. Details of these results are shown in Figure 4 and Table 3.

Table 3 Meta-analytic comparison of anxiety and depression interventions

		Effect		Heterogeneity		
		g	95% CI [a]	I^2	p	I^2 95 CI
Depression	All studies (n = 13)	1.178	1.06 - 1.29	95	<0.000	95 (94 - 97)
	Pre-post-correlation = .00	1.236	1.10 - 1.38	86		86 (78 - 91)
	Pre-post-correlation = .75	1.155	1.04 - 1.27	96		96 (94 - 97)
	Pre-post-correlation = .99	0.749	0.16 - 0.88	100		100 (100 - 100)
	Outliers Excluded [b]	1.176	1.09 - 1.26	75	0.001	75 (42 - 86)
	Without mixed treatments	1.282	1.26 - 1.44	89	<0.000	89 (84 - 92)
Anxiety	All studies (n = 20)	0.94	0.83 - 1.06	74	<0.000	74 (60 - 83)
	Pre-post-correlation = .00	0.95	0.83 - 1.07	93	<0.000	93 (91 - 95)
	Pre-post-correlation = .75	0.93	0.82 - 1.04	100	<0.000	100 (99 - 100)
	Pre-post-correlation = .99	0.70	0.62 - 0.78	77	<0.000	77 (62 - 86)
	Outliers Excluded [c]	0.90	0.81 - 0.99	90	<0.000	91 (88 - 95)
	Without mixed treatments	0.95	0.81 - 1.10	83	<0.000	83 (74 - 89)
	Without OCD treatments	0.93	0.81 - 1.05	84	<0.000	84 (76 - 90)
	Without PTSD treatments	0.88	0.78 - 0.98	95	<0.000	95 (94 - 97)
	Neither OCD nor PTSD	0.87	0.77 - 0.98	86	<0.000	86 (78 - 91)

[a] Confidence Interval

[b] Three excluded studies [45–47] as well as Ruwaard and colleagues depression study [48]

[c] Two excluded studies [49,50] as well as Ruwaard and colleagues PTSD and panic disorder studies

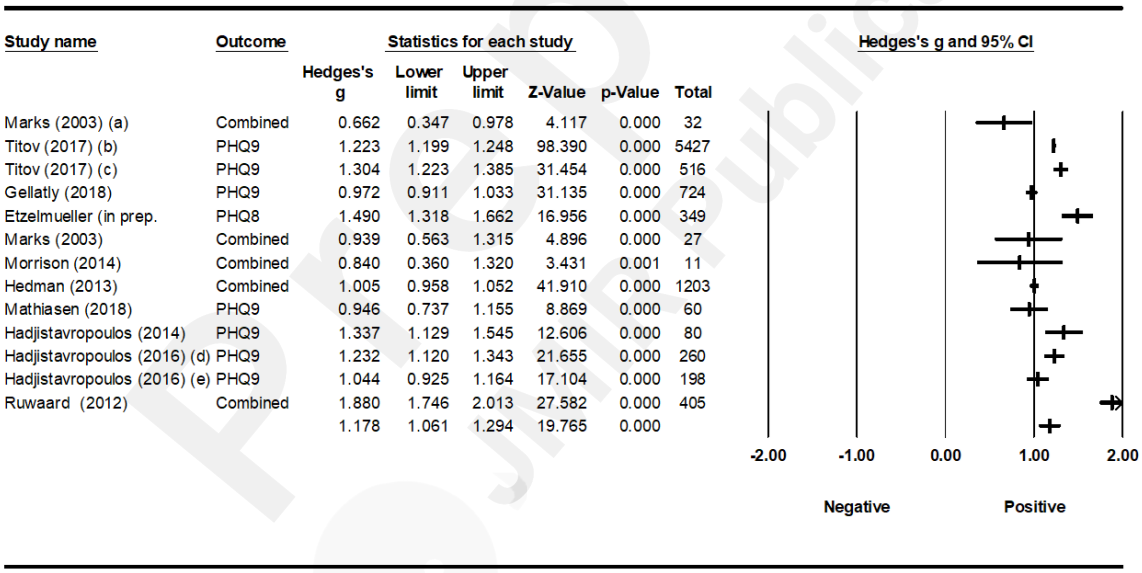
[48]and Titov and colleagues PTSD study [51].

In the main analysis described above, the pre-post measurement correlation was set to 0.59. Sensitivity analysis with correlations set to 0, 0.75 and 0.99 resulted in comparable effect sizes ($g_{Corr=0}$ = 0.95, $I^2_{Corr=0}$ = 74, 95% CI 60 - 83; P < .001; $g_{Corr=.75}$ = 0.93, $I^2_{Corr=.75}$ = 93, 95% CI 91 - 95; P < .001) with $g_{Corr=0.99}$ = 0.70 ($I^2_{Corr=.99}$=99, 95% CI 99 - 100; P < .001) resulting in the smallest effect size.

Both the visual inspection of the funnel plot and Egger`s test (P = .91) did not indicate a potential publication bias.

We found five studies to be outliers, since their results did not overlap with the 95% CI of the pooled estimate. Removing studies [49,50], as well as Ruwaard and colleagues PTSD and

Figure 3 Standardized effect sizes of iCBT treatments for depression in routine care (all studies included) [a]

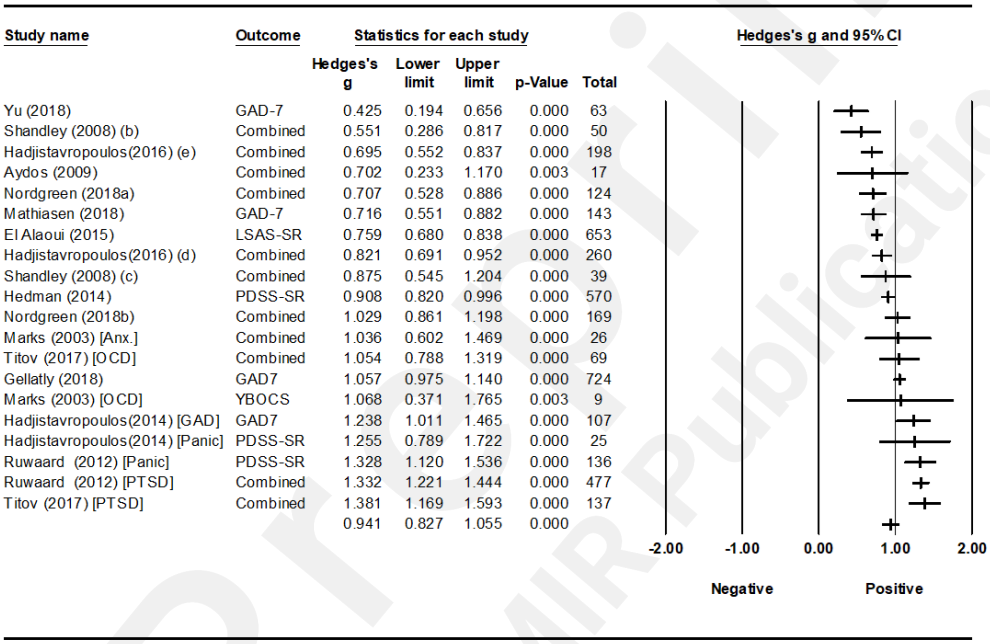


[a] Mixed depression and anxiety treatment
[b] Depression treatment
[c] Depression treatment for older adults
[d] Specialised Care
[e] Non-Specialised Care

panic disorder studies [48], and Titov and colleagues PTSD study [51] from the analysis did not influence the result significantly (g = 0.90; 95% CI = 0.81 – 0.99; I^2 = 77%, 95% CI = 62 - 86; p < .001), but resulted in less, while still high, heterogeneity. Excluding the mixed anxiety and depression studies did not result in a significantly different effect-size (g = 0.99;

95% CI = 0.86 – 1.12; I^2 = 92%; 95% CI = 88 - 95; P < .001). Neither removing OCD treatments (g = 0.93; 95% CI = 0.82 – 1.05; I^2 = 90%; 95% CI = 86-93; P < .001), PTSD treatments (g = 0.88; 95% CI = 0.78 – 0.98; I^2 = 83%, 95% CI = 74-89; P < .001) or both (g = 0.87; 95% CI = 0.77 – 0.98; I^2 = 84%; 95% CI = 76-90; P < .001) resulted in significantly different effect sizes, while lowering the heterogeneity.

Figure 4 Standardized effect sizes of iCBT treatments for anxiety in routine care (all studies included) [a]



[a] Full references are available in Appendix D. Marks Marks' study [47]is not providing an anxiety measure for the mixed depression and anxiety treatment, therefore this study has not been included in the analysis. "Combined" indicates that multiple measures for the main outcome have been combined in the analysis
[b] General Practitioner guided
[c] Therapist guided
[d] Specialised Care
[e] Non-Specialised Care

iCBT negative effects.

Appendix H comprises the results on negative effects. Less than half of the studies reported on deterioration rates (k = 7; 41%), with an average deterioration rate 2.9% (k = 14, 95% CI 1.9% - 4.3%; depression: k = 5, M = 2.5%, 95% CI 2.2% - 2.9%; anxiety: k = 9, M = 3.1%, 95% CI = 1.6% - 5.9%; the forest plot can be retrieved from the corresponding author).

No study reported other negative effects, one study mentioned that there were “no adverse outcomes”. No studies reported predictors of deterioration or other negative effects.

Subgroup analysis for of iCBT for the treatment of depression, anxiety or mixed depression and/ or anxiety.

Tables 4 and 5 show the results of all examined subgroup analyses. Significant differences between subgroups were found for professional training of coaches, supervision of coaches, and treatment duration for both depression and anxiety studies, and for recruitment pathways for depression studies only. Studies evaluating a period of nine to 13 weeks of treatment duration reported a significant lower effect size (depression: g = 1.00, 95% CI = 0.95 - 1.05; I^2 = 0; 95% CI = 0 - 85; anxiety: g = 0.83, 95% CI = 0.72 – 0.95; I^2 = 59; 95% CI = 9 - 81) compared to studies with less than 9 (depression: g = 1.17, 95% CI = 1.01 – 1.32; I^2 = 95; 95% CI = 92 - 97; anxiety: g =1.16, 95% CI = 0.97 – 1.34; I^2 = 93; 95% CI = 61 -91) or more than 13 weeks (depression: g = 1.37, 95% CI = 1.00 -1.74; I^2= 97; 95% CI = 94 - 98; anxiety: g = 0.98, 95% CI = 0.78 -1.17; I^2=89; 95% CI = 78 - 94). However, effect sizes within all examined subgroups were high.

Depression studies that recruited in community settings only reported significantly higher effect sizes (g = 1.37, 95% CI = 1.16 – 1.59; I^2 = 96; 95% CI = 94 - 98), compared to studies that recruited in clinical or clinical and community setting (g = 1.05, 95% CI = 0.95 – 1.14; I^2 = 78; 95% CI = 56 - 89). Across all recruitment pathways effect sizes were large, but heterogeneity remained high. We did not find this difference in the anxiety studies.

Studies only involving guides, not trained in CBT, showed a significantly lower effect size in depression studies ($g_{\text{Non-professional, Depression}}$ = 0.92, 95% CI = 0.79 – 1.05; I^2= 22; 95% CI =.0 - 88) than all other studies, including specifically trained professionals ($g_{\text{Other,Depression}}$ = 1.27, 95% CI = 1.14 -1.40; I^2 = 96; 95% CI = 94 -97). We did not find this effect in anxiety studies ($g_{\text{Non-professional,Anxiety}}$ = 0.87, 95% CI = 0.17 – 1.27; I^2 = 88; 95% CI = 73 - 95; $g_{\text{Other,Anxiety}}$ = 0.96, 95% CI = 0.83 -1.09; I^2 = 90; 95% CI = 85 - 93).

Depression studies reporting to having provided supervision to their coaches, trained their professionals and provided an intervention manual reported a significantly higher effect size ($g_{\text{Supervision}}$ = 1.27, 95% CI = 1.13 – 1.41; I^2 = 96; 95% CI = 94 - 97; g_{Training} = 1.35, 95% CI = 1.20 -1.51; I^2 = 95; 95% CI = 91 - 97; g_{Manual} = 1.47, 95% CI = 1.23 – 1.71; I^2 = 97; 95% CI = 95 - 98) compared to studies not reporting to provide these ($g_{\text{NoSupervision}}$ = 0.91, 95% CI = 0.75 – 1.08; I^2 = 39; 95% CI = 0 - 79; $g_{\text{NoTraining}}$ = 0.98, 95% CI = 0.94 – 1.02; I^2 = 7; 95% CI = 0 - 76;

$g_{NoManual}$ = 1.04, 95% CI = 0.95 – 1.13; I^2 = 75; 95% CI = 51 - 87). For anxiety studies we found similar effects for the reporting of training and providing an intervention manual ($g_{Training}$ = 1.07, 95% CI = 0.89 – 1.26; I^2 = 90; 95% CI = 85 - 94; g_{Manual} = 1.30, 95% CI = 1.19 – 1.41; I^2 = 29; 95% CI = 0 - 74 compared to $g_{NoTraining}$ = 0.80, 95% CI = 0.67 – 0.93; I^2 = 83; 95% CI = 70 - 88; $g_{NoManual}$ = 0.88, 95% CI = 0.75 – 0.94; I^2 = 81; 95% CI = 70 - 88), but not for supervision.

There were no differences between subgroups regarding all other examined subgroups, both for depression and anxiety studies.

Table 4 Subgroup-analyses: Depression treatments

		Effect		Heterogeneity			Subgroup analysis	
Subgroup Analysis [a]	N [b]	g	95% CI	I^2	p	I^2 95% CI [c]	Q-value	p(Q)
Recruitment pathway [d]								
Clinical and Community + Clinical	8	1.05	0.95 - 1.14	78	<0.001	56-89	7.253	0.007
Community	5	1.38	1.16 - 1.59	96	<0.001	94-98		
Specific treatment								
Mixed treatment	7	1.10	0.98 - 1.22	93	<0.001	87-96	0.736	0.391
Disorder specific treatment	6	1.27	0.91 - 1.62	97	<0.001	95-98		
Diagnosis conducted [e]								
Interview	7	1.127	0.89 - 1.35	97	<0.001	95-98	0.946	0.331
Questionnaire	5	1.25	1.16 - 1.34	81	<0.001	57-92		
Clinical cut-off / minimal symptom severity								
Yes	4	1.27	1.08 - 1.45	84	<0.001	60-94	0.897	0.344
No	7	1.17	0.98 - 1.35	96	0.001	94-97		
Treatment duration								
< 9 weeks	5	1.17	1.01 - 1.32	95	<0.001	92-97	7.485	0.024
9-13 weeks	4	1.00	0.95 - 1.05	0	0.85	0-85		
> 13 weeks	4	1.37	1.00 - 1.74	97	<0.001	94-98		
Guide CBT training (profession) [f]								
Non-professional	4	0.92	0.79 - 1.05	22	0.28	0-88	14.151	<0.000
Other	9	1.27	1.14 - 1.40	96	<0.001	94-97		
Guide supervision provided								
No	4	0.91	0.75 - 1.08	39	0.18	0-79	10.339	0.001
Yes	9	1.27	1.13 - 1.41	96	<0.001	94-97		
Guide training provided								
No	6	0.98	0.94 - 1.02	7	0.37	0-76	21.368	<0.000
Yes	7	1.35	1.20 - 1.51	95	<0.001	91-97		
Intervention manual provided								
No	9	1.039	0.95 - 1.137	75	<0.001	51-87	10.715	0.001
...Yes	4	1.467	1.23 - 1.71	97	<0.001	95-98		

		Effect		Heterogeneity			Subgroup analysis	
Risk of Bias - Researcher Allegiance								
High	7	1.252	1.08 -1.42	97	<0.001	95-98	1.347	0.246
Low	5	1.119	0.99 - 1.29	70	0.010	23-88		
Risk of Bias - Confounding (Treatment Inclusion)								
High	7	1.215	1.06 -1.43	96	<0.001	94-98	0.985	0.321
Low	5	1.105	0.98 - 1.23	82	<0.001	58-92		

[a] Test against "Guidance format: face-to-face vs written guidance", "Guidance modality: Message, Email, Telephone, F2F" and "Guide profession" excluded, as there were too few studies included in analysis

[b] Number of studies

[c] Confidence Interval

[d] Only two studies included via the clinical pathway only. We combined the categories "Both, community and clinical" and "clinical" for this analysis

[e] Excluding one study [52], as this is the only study using clinical judgement without specifying the use of an interview or questionnaire

[f] We grouped all studies involving guides not specifically trained in delivering CBT in the category “non-professional”, and studies involving psychiatrists, psychologists, or psychotherapists in their guidance in the category “other”.

Table 5 Subgroup-analyses: Anxiety treatments

		Effect		Heterogeneity			Subgroup analysis	
	N [a]	g	95% CI	I^2	p	I^2 95% CI [b]	Q-value	p(Q)
Recruitment pathway								
Clinical	5	0.77	0.53 - 1.01	91	<0.001	81-95		
Community	7	1.08	0.85 - 1.31	88	<0.001	78-94	3.340	0.188
Community + Clinical	8	0.90	0.78 - 1.01	74	<0.001	46-87		
Specific Disorder								
Panic	6	0.95	0.71 - 1.13	91	<0.001	64-92		
Non-panic treatments	14	0.92	0.801 - 1.09	83	<0.001	(-)	0.053	0.818
Guidance: Modality					<0.001			
Email	7	1.11	0.95 - 1.26	56	<0.001	0-81		
Message	8	0.88	0.69 - 1.06	94	<0.001	90-96	4.744	0.093
Synchronous (Tel. or F2F)	5	0.86	0.66 - 1.10	83	<0.001	60-92		
Guide CBT training (profession) [c]								
Non-professional	4	0.87	0.47 -1.27	88	<0.001	73-95	0.165	0.685
Other	16	0.96	0.83 - 1.09	90	<0.001	85-93		
Guidance: Moment								
Weekly/Bi/weekly	10	0.66	0.73 - 1.00	74	<0.001	0-85	0.174	0.677
Reaction	4	0.83	0.72 - 0.94	53	<0.001	50-86		
Guide supervision provided								
No	6	0.82	0.70 - 0.94	57	0.041	0-83	2.812	0.094
Yes	14	0.9811	0.83 - 1.13	90	<0.001	86-94		
Guide training provided								

		Effect		Heterogeneity			Subgroup analysis	
No	10	0.80	0.67 - 0.93	83	<0.001	70-90	5.779	0.016
Yes	10	1.07	0.89 - 1.26	90	<0.001	85-94		
Intervention manual provided								
No	16	0.88	0.75 - 0.94	81	<0.001	70-88	37.209	<0.000
...Yes	4	1.30	1.19 - 1.41	29	0.241	0-74		
Approach to data analysis								
Completer	4	1.05	0.98 - 1.12	0	<0.001	0-77	2.796	0.096
ITT	16	0.92	0.78 - 1.06	91	<0.001	86-94		
Diagnostic method								
Interview	15	0.97	0.84 - 1.06	88	<0.001	83-92	0.388	0.533
Questionnaire	5	0.87	0.6 - 1.14	91	<0.001	82-95		
	N [a]	g	95% CI	I^2	p	I^2 95% CI [b]	Q-value	p(Q)
Treatment duration								
< 9 weeks	5	1.16	0.97 - 1.34	83	<0.001	61-91	8.686	0.013
9-13 weeks	8	0.83	0.72 - 0.95	59	0.018	9-81		
> 13 weeks	6	0.98	0.78 - 1.17	89	<0.001	78-94		
Risk of Bias - Researcher Allegiance								
High		0.99	0.83 - 1.14	90	<0.001	0-60-	1.613	0.204
Low		0.82	0.63 - 1.02	39	<0.001	66-92		
Risk of Bias - Confounding (Treatment Inclusion)								
High		1.03	0.89 - 1.18	89	<0.001	83-93	4.852	0.028
Low		0.82	0.70 - 0.93	69	<0.001	37-84		

[a] Number of studies

[b] Confidence Interval

[c] We grouped all studies involving guides not specifically trained in delivering CBT in the category “non-professional”, and studies involving psychiatrists, psychologists, or psychotherapists in their guidance in the category “other”

Subgroup analyses comparing studies rated with high versus low risk indicated that R*esearcher Allegiance* did not have a significant influence on the estimated effect sizes for neither anxiety nor depression studies. The heterogeneity within the studies reporting low risk of bias on R*esearcher Allegiance* did reveal an I^2 of 39 compared to an I^2 of 89.63 for studies reporting a high risk of bias. Moreover, anxiety studies being rated as at high risk of *Treatment Inclusion confounding* had higher estimated effect sizes. This was not replicated in subgroup analyses of interventions targeting depression. Anxiety studies at high risk of *Selection Bias* reported significantly lower effect sizes. Similar outcomes were not replicated in depression trials.

Meta-regression analysis for iCBT for the treatment of anxiety, depression or mixed depression and anxiety.

Meta-regression analyses indicated that longer treatment duration in depression studies was positively associated with a higher effect (P = .02; *est*. = 0.03, R^2 = 0.00). This effect was not found in anxiety studies (P = .94). No other of the examined variables, ie. guidance time, number of contacts, number of sessions completed, or the percentage of treatment completers, was significantly associated with the observed effect sizes, neither in depression nor anxiety studies.

Discussion

The present study aimed to examine the acceptability, effects on symptom change and negative effects of guided iCBT interventions in treating depression and anxiety in routine care. Regarding the uptake of the service, on average 70.2% of people screened were not offered inclusion, and of those included, 73.0% started the intervention. The vast majority of participants reached were female with an average age of 38.3 years, and 61.3% of participants completed the interventions as planned. Reported participant satisfaction was high, although inconstant reported results did not allow us to pool effects. The average professional guidance time per participant was 133.49 minutes over the treatment duration. With regard to effects on symptom change, results indicated large average reductions for both depression (g = 1.18, 95% CI = 1.06 – 1.29) and anxiety (g = 0.94, 95% CI = 0.83 – 1.062). However, the heterogeneity between studies was high. Nevertheless, all examined effect

sizes were at least moderate indicating the intervention's potential when delivered under routine care conditions with effects ranging from moderate to large. Average deterioration rates were 3.2% for depression and 3.1% for anxiety. Subgroup analyses indicated a range of iCBT service-related characteristics to be associated with the observed treatment effects.

Regarding uptake, we found that many participants getting in contact with the iCBT service did not start the intervention. Pre-treatment drop-out is hard to assess and accordingly, reasons for not starting an iCBT intervention after inclusion have not been discussed in the original publications.

The average age of participants found in this study (M = 38.30) appears to be slightly lower than reported in RCTs on guided iCBT interventions for the treatment of depression (M = 42.5 [53]) but comparable to reports on the mean age of participants within guided iCBT interventions for the treatment of anxiety [54]. The percentage of females in the routine care study population was higher for depression studies compared to guided iCBT for the treatment of depression [53] and similar to reports on participants in guided iCBT interventions for the treatment of anxiety [54] in experimental settings. As similar distributions between female and male users are reported in face-to-face mental health service utilization [55], this effect might rather be explained by gender differences in help-seeking behaviour than being related to iCBT service-related factors [56] as well as by gender differences in the prevalence of depression and anxiety disorder [57,58]. Future studies should focus on ways to attract men to use iCBT interventions.

The pooled reported percentage of sessions completed of 62.6% in depression and 57.3% in anxiety studies was lower than described in meta-analyses on adherence in RCTs on iCBT interventions. Comparing the adherence to iCBT and face-to-face CBT, van Ballegooijen and colleagues [59] reported that on average participants completed 80.8% of treatment sessions in the iCBT and 83.9% in the face-to-face intervention [59]. Similarly, the percentage of participants completing the treatment as planned was lower (62.8% for depression and 61.7% for anxiety studies) than reported elsewhere [59,60]. These differences might be due to the assumed adherence-fostering effect of randomised-controlled settings versus routine care [61]. However, completion rates were reported inconsistently across studies, applying different criteria such as study- or treatment completer including several definitions of treatment completions. To facilitate comparability, literature on iCBT completion should settle on one reporting standard. Further investigating factors promoting the acceptance of

iCBT interventions, also when reporting on effectiveness results in routine care, may lead to a deeper understanding which then might foster intervention development and upscaling.

Results on the effectiveness of iCBT ($g_{Depression}$ = 1.18, 95% CI = 1.06 – 1.29 and $g_{Anxiety}$ = 0.94, 95% CI = 0.83 – 1.062) confirm findings of recently published systematic reviews and meta-analyses on randomised control trials of iCBT for depression and anxiety. Königbauer and colleagues [12] found medium to large pre-post within-group effects ranging between -0.64 and -2.24 for interventions treating clinical depression [12]. To our knowledge, no recent meta-analysis reported on pre-post effect sizes of studies targeting guided iCBT interventions for the treatment of anxiety. On an individual study level, pre-post effects in randomized trials range between 0.54 to 2.40 (please see Appendix I for references) compared to 0.66 to 1.88 in depression and 0.42 to 1.38 (*Hedges' g*) in anxiety within this analysis.

With regard to randomized pragmatic trials conducted under routine care conditions, Andrews and colleagues [15] examined a sample of 64 papers reporting results of randomized controlled trials on the effectiveness of iCBT for the treatment of depression, panic disorder, generalized anxiety disorder and social phobia in comparison to control groups in routine practice. This review study reported effect sizes for depression, panic disorder, generalised anxiety disorder and social phobia ranging between g = 0.67 - 1.31 [15]. The same study identified eight papers investigating the effectiveness of iCBT reporting an average effect size of g = 1.07 across the treatment of depression, panic disorder, generalized anxiety disorder and social phobia [15]. Between group effects were moderate to large (g = 0.72; 95% CI 0.60-0.83; p < .001; of I^2 = 53%, 95% CI 31-66) in the most recent meta-analysis of iCBT treatments for anxiety compared to control conditions in reducing symptoms of anxiety in an adult population [13]. Additionally, the present results are in line with meta-analytic findings on face-to-face CBT treatments implemented in routine care with pre-post effect-size found in randomized trials ranging from d = 0.69 to 2.28 for depression [27] and g = 0.73 to 2.59 for anxiety treatments [62].

Results on deterioration rates (3.2% in depression and 3.1% in anxiety studies) were slightly lower, but within the 95% CI of findings based on randomized controlled trials for internet-based guided self-help interventions (3.36%) for depression [63] and anxiety (5.8% [64]), and also comparable to deterioration rates in face-to-face psychotherapy for depression [65]. Criteria defining deterioration varied between studies, and unfortunately,

neither were reports on other negative effects included in most primary studies, nor reported any study predictors of deterioration. This seems of utmost importance in order to identify those individuals that should potentially be referred to other mental health services Their investigation is of specific importance within naturalistic study designs and under routine care conditions [63,64,66].

Most evaluated iCBT services for depression (69.2%) excluded severe cases and individuals with suicidal ideation (*k* = 9/13) at baseline. However, a large-scale study showed that iCBT services can also result in positive effects on suicidality, reducing the prevalence of suicidal ideation from 50% at baseline to 27% after treatment [67]. Also, a recent individual patient data meta-analysis on randomized controlled trials indicated guided iCBT also to result in clinically meaningful results in individuals with severe depression symptomatology [53]. Given that many individuals applying to iCBT services either do not have access to other immediate care or are not willing to utilize alternative treatment services, future studies should explore the balance between potential risk and benefits of opening up those services also to populations showing elevated suicidal ideation. In such cases, it seems of utmost importance to monitor potential upcoming crises using standard operating procedures involving trained clinicians and to evaluate treatment success at the end of the service. In case of non-response, individuals should be motivated and guided to utilize other mental health care services, if available. Such standardized crisis procedures were only reported to be employed by less than half of the studies included in the present review. ICBT services in routine care might profit from clear pathways of referral to other services in case of non-response and symptom deterioration. Furthermore, future research should facilitate our understanding of the effects of routine outcome monitoring in in routinely applied iCBT [68], as this monitoring could help evaluate participant's progress throughout the course of treatment, using standardized outcome measures to elicit client as part of a measurement based care delivery approach in routine mental and behavioral health care [69,70].

The finding that treatment outcomes of depression interventions were greater when recruitment was carried out using an open recruitment strategy in a community setting compared to when recruited in a clinical setting, is in line with finding from Romijn and colleagues [13] with regard to randomized pragmatic studies on anxiety disorder treatments. However, in our study, this interaction was only found for depression and could not be confirmed for anxiety disorders. One potential explanation for the difference in effects might

be differences in the characteristics of included patients. There is evidence that iCBT recruiting via open recruitment strategies such as through web-based channels only might reach a specific population that is different from those seeking help in a clinical setting [19]. It is often argued that internet-interventions might reach individuals that would otherwise not seek treatment or only at a later time point. Given that for example chronicity of depression is associated with worse treatment outcomes [71], the difference in effect might be explained by reaching a population with lower chronicity. However, such an assumption needs to be confirmed in future studies.

Further subgroup analyses indicated iCBT services for the treatment of depression utilizing trained professionals (psychotherapists, psychiatrists) to result in larger pre-post changes, compared to iCBT services that used only non-professionals not trained in CBT (psychologists without specialized CBT training, nurses, GPs, counselors, coaches, lived experience coordinators). However, we did not find this effect within the anxiety studies. Also, effects in the subgroup of depression studies involving non-professionals were large, indicating the potential to deliver iCBT services e.g. in contexts when there might be a shortage of trained clinicians. In case non-professionals deliver guidance in iCBT services, supervision by trained clinicians, including the availability of professionals for crisis intervention seems warranted. Further subgroup analyses also indicated that providing supervision to coaches is also associated with higher average treatment effects, for depression, but not for anxiety studies. Furthermore, training the professional and providing an intervention manual is positively related to the interventions' effectiveness. This result must be interpreted with caution as we coded all studies not mentioning supervision, training or manual provision in their publication as not providing these components. Furthermore, these components do not inform us about actual treatment fidelity. Further research should focus on the effects on treatment outcomes of providing supervision, training and intervention manuals to professionals working with iCBT interventions in routine care as well as the assessment of treatment fidelity.

Moreover, we did not find a difference in effects on mean symptom change between iCBT services who applied diagnostic interviews for patient allocation vs those that used self-reports only. This is in line both, with meta-analytic findings from randomized controlled trials on guided digital interventions for depression [53] and with studies directly comparing the effectiveness of iCBT services when treatment allocation was based on an automatic

web-based assessment vs. clinician assessment [72]. This indicates that such services can be used in contexts when implementing services with initial clinician-assessment is not possible, without effecting average treatment success. However, it must be noted, that although results might not indicate differences on the group level, it might be the case that using web-based assessments only, without a clinical assessment, will overlook relevant diagnostic information that require immediate attention, such as suicidal risk or an underlying treatment need for comorbid disorders such as PTSD on an individual level.

Strengths of this meta-analysis include the exclusive focus on evaluating iCBT interventions for their acceptability and clinical outcomes under real-world conditions. Unlike previous systematic reviews that mixed efficacy with effectiveness trials, in this review we focused only on studies conducted in regular care settings. This is important as we strive for reporting routine care results free from biases possibly being introduced within efficacy studies such as a stricter application of protocolised procedures, eligibility criteria and randomisation [19–22]. Moreover, we presented an overview of implementation indicators existing in the included studies that can be used to gain a better understanding of how iCBT can be adopted by regular care services. Nevertheless, the current findings should be interpreted with caution due to several limitations.

First, the heterogeneity in our sample was high and significant illustrating a great variation in the results of the included studies. Thus, we cannot draw firm conclusions regarding the average effect of iCBT in routine care. Moreover, within group effect sizes do not depict an optimal estimator for the treatment effect because they are not independent of each other and do not account for recovery occurring independent of the treatment, thereby leading to an overestimation of the treatment effect [40]. However, in comparison with and on the basis of the reported efficacy of iCBT interventions established in RCTs, they depict the best available indicator of the effects of iCBT solutions in a routine care environment. Furthermore, we found that treatment duration had a significant influence on treatment effects. This result also supports the hypothesis that findings on pre post changes in symptom severity might have been influenced by spontaneous or unexplained recovery, which is a common factor in depression [73]. However, our main results are in line with within-group effect sizes found in RCTs, where the spontaneous recovery also occurs, and we, therefore, conclude that our effects can be considered substantial. Although heterogeneity was not explained by any other of the examined subgroups, several assumptions can be

made regarding its sources. One other explanation for the high heterogeneity might be the influence of contextual factors of observational studies, such as sampling methods, participant characteristics, within group effect sizes, and differences between the studies in reporting outcomes. It can be hypothesised that a greater harmonization regarding the conduct and reporting of effectiveness studies in routine care could lead to greater comparability of the studies' results. Another reason for the observed heterogeneity might be the different contexts of the regular care facilities across different countries. There is a great variability in the degree of e-mental health penetration in different countries. For instance, Australia is considered one of the frontrunners in the e-mental health field, while Norway adopted these interventions very recently [74]. Thus, professionals might differ in the way they interact with e-mental health around the world. Finally, the interventions might differ in the way that they have been developed. These results also imply the importance of establishing a firm evidence base for individual iCBT interventions before their larger upscale.

Second, firm conclusions on treatment effects might be biased by studies also including participants which could also participate in other psycho-therapeutic treatments. Meanwhile, the data does not allow conclusions on the percentage of participants receiving additional treatment and represents the routine practice. Additionally, no study reported adjusting for confounders such as baseline symptom-severity, treatment fidelity (provision and use) or changes in the treatment over the course of the studies, which should be considered in future reports on effects of iCBT in routine care.

Future studies should add to the body of literature on iCBT interventions examined under routine care conditions. Additionally, these studies should not solely focus on the effectiveness of the interventions, but if possible, it would be helpful if they also reported on specific service-, implementation- and context-related outcomes. One way of achieving this might be through taxonomy and guidelines for the reporting of iCBT effectiveness, implementation and context outcomes in routine care. In contrast to standards of reporting randomized controlled trials, no such international standards exist when it comes to reporting non-randomized intervention studies. Proctor and colleagues [75] suggested a list of outcomes for implementation related research and Hermes and colleagues [76] recently made suggestions on how to build upon these ideas to establish a measurement system for the implementation of behavioural intervention technologies. Moreover, such research should always be discussed and evaluated in the light of quality criteria established to help

all involved stakeholders, patients, practitioners and decision makers on local- and policy-level to identify not only effective but also safe interventions [77].

In conclusion, this study provides further evidence supporting the acceptability and effectiveness of guided iCBT for the treatment of depression and anxiety when implemented into routine care, while results on negative effects are less clear. Guided iCBT may be an effective way of overcoming barriers to treatment provision. It may substantially increase the coverage of usual care services and offer an innovative treatment format in the treatment of depression and anxiety.

Authors Contributions

AE and DDE initiated and conceptualized the study and drafted the manuscript with initial feedback from all authors. AE, CV and DEE acquired and managed the data. AE, CV and EK conducted the data analysis, with critical feedback from DDE and PC. All authors revised and edited the manuscript, provided final approval to this version and agreed to be accountable for this work.

Conflict of Interest

Assoc. Prof. Ebert reports to have received consultancy fees or served in the scientific advisory board from several companies such as Minddistrict, Sanofi, Lantern, Schön Kliniken, German health insurance companies (BARMER and Techniker Krankenkasse), and chambers of psychotherapists. Dr. Ebert is one of the stakeholders of the Institute for health trainings online (GET.ON), which aims to implement scientific findings related to digital health interventions into routine care. Anne Etzelmueller is employed by the Institute for health trainings online (GET.ON) as research coordinator. Prof. Titov is funded by the Australian Government to develop and provide a free national online and telephone-delivered treatment service. Prof. Baumeister served in the e-mental-health associated scientific advisory boards, e-mental health interest groups and task forces. All other authors do not report a conflict of interest.

Acknowledgements

The authors want to acknowledge M.Sc. Psych. Silja Schenk's assistance in managing the literature search in an early stage of the project. We also want to thank Dr. Robin Kok for providing his knowledge on his website https://www.robinkok.eu/.

Multimedia Appendices

Multimedia Appendix A. Prisma Checklist

Multimedia Appendix B. Search strategy

Multimedia Appendix C. Risk of Bias Assessment Definition

Multimedia Appendix D. References of included studies

Multimedia Appendix E. iCBT service-related characteristics (I)

Multimedia Appendix F. iCBT service-related characteristics (II)

Multimedia Appendix G. Acceptability – Uptake and participant characteristics

Multimedia Appendix H. Acceptability - Participant satisfaction and negative effects

Multimedia Appendix I. References for studies targeting guided iCBT interventions for the treatment of anxiety reporting on pre-post effect sizes

References

1. Kessler RC, Bromet E. The epidemiology of depression across cultures. Annu Rev Public Health. 2013;34:119–138. PMID: 23514317
2. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The Epidemiology of Major Depressive Disorder Results From the National Comorbidity Survey Replication (NCS-R). Health Care (Don Mills). 2003;289(23):3095–3105.
3. Donohue J, Pincus H. Reducing the Societal Burden of Depression. Pharmacoeconomics. Springer International Publishing; 2007;25(1):7–24.
4. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet. Elsevier; 2013;382(9904):1575–1586.
5. Cuijpers P, Karyotaki E, de Wit L, Ebert DD. The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. Psychother Res. 2019;1–15.
6. Wittchen H-U, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. Elsevier; 2011;21(9):655–79. PMID: 21896369
7. Mack S, Jacobi F, Gerschler A, Strehle J, Höfler M, Busch MA, et al. Self-reported utilization of mental health services in the adult German population - evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). Int J Methods Psychiatr Res. 2014; PMID: 24687693
8. Ebert DD, Nobis S, Lehr D, Baumeister H, Riper H, Auerbach RP, et al. The 6-month effectiveness of Internet-based guided self-help for depression in adults with Type 1 and 2 diabetes mellitus. Diabet Med. 2017;34(1):99–107.
9. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, et al. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health: A Report of the EFPA E-Health Taskforce [Internet]. Eur Psychol. Hogrefe Publishing; 2018. p. 167–187.
10. Ebert DD, Cuijpers P, Muñoz RF, Baumeister H. Prevention of Mental Health Disorders Using Internet- and Mobile-Based Interventions: A Narrative Review and Recommendations for Future Research. Front Psychiatry. Frontiers; 2017;8(AUG):116.
11. Sander L, Rausch L, Baumeister H. Effectiveness of Internet-Based Interventions for the Prevention of Mental Disorders: A Systematic Review and Meta-Analysis. JMIR Ment Heal. 2016;3(3):e38. PMID: 27535468
12. Königbauer J, Letsch J, Doebler P, Ebert DD, Baumeister H. Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. J Affect Disord. 2017;223(April):28–40. PMID: 28715726
13. Romijn G, Batelaan N, Kok R, Koning J, van Balkom A, Titov N, et al. Internet-Delivered

Cognitive Behavioral Therapy for Anxiety Disorders in Open Community Versus Clinical Service Recruitment: Meta-Analysis. J Med Internet Res. 2019;21(4):e11706.

14. Andersson G, Carlbring P, Titov N, Lindefors N. Internet Interventions for Adults with Anxiety and Mood Disorders: A Narrative Umbrella Review of Recent Meta-Analyses. Can J Psychiatry. SAGE PublicationsSage CA: Los Angeles, CA; 2019;070674371983938.

15. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. J Anxiety Disord. Elsevier Ltd; 2018. p. 70–78. PMID: 29422409

16. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis [Internet]. Cogn Behav Ther. Routledge; 2018. p. 1–18.

17. Cuijpers P, Noma H, Karyotaki E, Cipriani A, Furukawa TA. Effectiveness and Acceptability of Cognitive Behavior Therapy Delivery Formats in Adults with Depression: A Network Meta-analysis. JAMA Psychiatry. American Medical Association; 2019;

18. Wells KB. Treatment Research at the Crossroads: The Scientific Interface of Clinical Trials and Effectiveness Research. Am J Psychiatry. American Psychiatric Publishing; 1999;156(1):5–10.

19. Rothwell PM. External validity of randomised controlled trials:"to whom do the results of this trial apply?" Lancet. 2005;365(9453):82–93.

20. Singal AG, Higgins PDR, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol. Nature Publishing Group; 2014;5(1):e45. PMID: 24384867

21. Ebert DD, Baumeister H. Internet-Based Self-help Interventions for Depression in Routine Care. JAMA Psychiatry. American Medical Association; 2017;74(8):852.

22. van der Lem R, van der Wee NJA, van Veen T, Zitman FG. Efficacy versus Effectiveness: A Direct Comparison of the Outcome of Treatment for Mild to Moderate Depression in Randomized Controlled Trials and Daily Practice. Psychother Psychosom. Karger Publishers; 2012;81(4):226–234. PMID: 22584117

23. Vlahakes GJ. The value of phase 4 clinical testing. N Engl J Med. 2006;354(4):413–415.

24. Hill TP. Conducting Phase IV clinical studies: A moral imperative? Ecancermedicalscience. 2012.

25. Andersson G, Hedman E. Effectiveness of guided internet-based cognitive behavior therapy in regular clinical settings. Verhaltenstherapie. 2013;23(3):140–148. PMID: 2013612983

26. Erbe D, Eichert H-C, Riper H, Ebert DD. Blending Face-to-Face and Internet-Based Interventions for the Treatment of Mental Disorders in Adults: Systematic Review. J Med Internet Res. Journal of Medical Internet Research; 2017;19(9):e306. PMID: 28916506

27. Hans E, Hiller W. Effectiveness of and Dropout From Outpatient Cognitive Behavioral

Therapy for Adult Unipolar Depression: A Meta-Analysis of Nonrandomized Effectiveness Studies. 2013;

28. Zarski A-C, Lehr D, Berking M, Riper H, Cuijpers P, Ebert DD. Adherence to Internet-Based Mobile-Supported Stress Management: A Pooled Analysis of Individual Participant Data From Three Randomized Controlled Trials. J Med Internet Res. Journal of Medical Internet Research; 2016;18(6):e146. PMID: 27357528

29. National Institute for Health Research. Glossary. https://www.nihr.ac.uk/glossary?letter=F&postcategory=-1 [Internet].

30. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol. 2010;10:67.

31. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. The Cochra. JPT H, S G, editors. Cochrane Libr. 2006. PMID: 21050962

32. Sterne JAC, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. British Medical Journal Publishing Group; 2016;355:i4919. PMID: 27733354

33. Higgins JPT, Altman DG. Assessing Risk of Bias in Included Studies. Cochrane Handb Syst Rev Interv Cochrane B Ser. Chichester, UK: John Wiley & Sons, Ltd; 2008. p. 187–241. PMID: 17432993

34. Munder T, Brütsch O, Leonhart R, Gerger H, Barth J. Researcher allegiance in psychotherapy outcome research: An overview of reviews. Clin Psychol Rev. Pergamon; 2013. p. 501–511.

35. Hedges L, Olkin I. Statistical methods for meta-analysis. Academic press.; 2014.

36. Cohen J. Statistical power analysis. New York: Behavioral Sciences; 1988.

37. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2019.

38. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Heal. 2019;22(4):153–160.

39. Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ. Empirical Assessment of Within-Arm Correlation Imputation in Trials of Continuous Outcomes. Methods Res Reports. 2012;12(EHC141):EF. PMID: 23326900

40. Cuijpers P, Weitz E, Cristea IA, Twisk J. Pre-post effect sizes should be avoided in meta-analyses. 2016;

41. Borenstein M, Higgins JPT, Hedges L V., Rothstein HR. Basics of meta-analysis: I 2 is not an absolute measure of heterogeneity. Res Synth Methods. 2017;8(1):5–18.

42. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343(7818).

43. Sterne JAC, Egger M. Regression Methods to Detect Publication and Other Bias in Meta-Analysis. Publ Bias Meta-Analysis Prev Assess Adjust. John Wiley & Sons, Ltd; 2006. p. 99–110.

44. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. Biometric Society; 2000;56(2):455–463.

45. Gellatly J, Chisnall L, Seccombe N, Ragan K, Lidbetter N, Cavanagh K. @Home eTherapy Service for People with Common Mental Health Problems: An Evaluation. Behav Cogn Psychother. 2018;46(1):115–120.

46. Hedman E, Ljótsson B, Kaldo V, Hesser H, El Alaoui S, Kraepelien M, et al. Effectiveness of Internet-based cognitive behaviour therapy for depression in routine psychiatric care. J Affect Disord. Elsevier; 2014;155(1):49–58.

47. Marks IM, Mataix-Cols D, Kenwright M, Cameron R, Hirsch S, Gega L. Pragmatic evaluation of computer-aided self-help for anxiety and depression. Br J Psychiatry. Cambridge University Press; 2003;183(01):57–65.

48. Ruwaard J, Lange A, Schrieken B, Dolan C V., Emmelkamp P. The Effectiveness of Online Cognitive Behavioral Treatment in Routine Clinical Practice. Botbol M, editor. PLoS One. Public Library of Science; 2012;7(7):e40089.

49. Shandley K, Austin DW, Klein B, Pier C, Schattner P, Pierce D, et al. Therapist-assisted, internet-based treatment for panic disorder: Can general practitioners achieve comparable patient outcomes to psychologists? J Med Internet Res. JMIR Publications Inc.; 2008;10(2):e14. PMID: 18487138

50. Yu JS, Szigethy E, Wallace M, Solano F, Oser M. Implementation of a Guided, Digital Cognitive Behavioral Program for Anxiety in Primary Care: Preliminary Findings of Engagement and Effectiveness. Telemed e-Health. Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2018;24(11):970–978. PMID: 29480752

51. Titov N, Dear BF, Staples LG, Bennett-Levy J, Klein B, Rapee RM, et al. The first 30 months of the MindSpot Clinic: Evaluation of a national e-mental health service against project objectives. Aust N Z J Psychiatry. SAGE PublicationsSage UK: London, England; 2017;51(12):1227–1239.

52. Morrison C, Walker G, Ruggeri K, Hacker Hughes J. An implementation pilot of the MindBalance web-based intervention for depression in three IAPT services. Cogn Behav Ther. Cambridge University Press; 2014;7:e15.

53. Karyotaki E, Ebert DD, Donkin L, Riper H, Twisk J, Burger S, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clin Psychol Rev. Pergamon; 2018;63:80–92.

54. Olthuis J V., Watt MC, Bailey K, Hayden JA, Stewart SH. Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. Cochrane Database Syst Rev. John Wiley and Sons Ltd; 2016.

55. Seidler ZE, Dawes AJ, Rice SM, Oliffe JL, Dhillon HM. The role of masculinity in men's help-seeking for depression: A systematic review. Clin Psychol Rev. Elsevier Inc.; 2016. p. 106–118.

56. Thompson AE, Anisimowicz Y, Miedema B, Hogg W, Wodchis WP, Aubrey-Bassler K.

The influence of gender and other patient characteristics on health care-seeking behaviour: A QUALICOPC study. BMC Fam Pract. BioMed Central Ltd.; 2016;17(1).

57. Seedat S, Scott KM, Angermeyer MC, Berglund P, Bromet EJ, Brugha TS, et al. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry. 2009;66(7):785–795.

58. Kuehner C. Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. Acta Psychiatr Scand. 2003;108(3):163–174.

59. Van Ballegooijen W, Cuijpers P, Van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: A meta-analysis. PLoS One. Public Library of Science; 2014;9(7).

60. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev. 2012;32(4):329–42. PMID: 22466510

61. Ebert DD, Baumeister H. Internet- und mobil-basierte Interventionen in der Psychotherapie: Ein Überblick. Psychotherapeutenjournal. 2016;1:22–31.

62. Stewart RE. Cognitive-Behavioral Therapy for Adult Anxiety Disorders in Clinical Practice: A Meta-Analysis of Effectiveness Studies. Artic J Consult Clin Psychol. 2009;

63. Ebert DD, Donkin L, Andersson G, Andrews G, Berger T, Carlbring P, et al. Does Internet-based guided-self-help for depression cause harm? An individual participant data meta-analysis on deterioration rates and its moderators in randomized controlled trials. Psychol Med. Cambridge University Press; 2016;46(13):2679–2693.

64. Rozental A, Magnusson K, Boettcher J, Andersson G, Carlbring P. For Better or Worse: An Individual Patient Data Meta-Analysis of Deterioration Among Participants Receiving Internet-Based Cognitive Behavior Therapy. J Consult Clin Psychol. 2017;85(2):160–177.

65. Cuijpers P, Reijnders M, Karyotaki E, de Wit L, Ebert DD. Negative effects of psychotherapies for adult depression: A meta-analysis of deterioration rates [Internet]. J Affect Disord. 2018. p. 138–145. PMID: 30005327

66. Rozental A, Castonguay L, Dimidjian S, Lambert M, Shafran R, Andersson G, et al. Negative effects in psychotherapy: commentary and recommendations for future research and clinical practice. 2018;

67. Mewton L, Andrews G. Cognitive behaviour therapy via the internet for depression: A useful strategy to reduce suicidal ideation. J Affect Disord. Elsevier B.V.; 2015;170:78–84. PMID: 25233243

68. Kendrick T, El-Gohary M, Stuart B, Gilbody S, Churchill R, Aiken L, et al. Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. Cochrane Database Syst Rev. John Wiley and Sons Ltd; 2016.

69. Lambert MJ, Harmon KL. The merits of implementing routine outcome monitoring in clinical practice. Clin Psychol Sci Pract. Blackwell Publishing Inc.; 2018;25(4):e12268.

70. Boswell JF, Kraus DR, Miller SD, Lambert MJ. Psychotherapy Research Implementing

routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. Psychother Res. 2015;25.

71. Cuijpers P, Andersson G, Donker T, Van Straten A. Psychological treatment of depression: Results of a series of meta-analyses. Nord J Psychiatry. 2011;65(6):354–364.

72. Mason EC, Andrews G. The use of automated assessments in internet-based CBT: The computer will be with you shortly. Internet Interv. Elsevier; 2014;1(4):216–224.

73. Cuijpers P. The Challenges of Improving Treatments for Depression. JAMA. American Medical Association; 2018;320(24):2529.

74. Titov N, Dear B, Nielssen O, Staples L, Hadjistavropoulos H, Nugent M, et al. ICBT in routine care: A descriptive analysis of successful clinics in five countries. Internet Interv. 2018. p. 108–115.

75. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Heal Ment Heal Serv Res. Springer US; 2011;38(2):65–76.

76. Hermes E DA, Lyon AR, Schueller SM, Glass JE. Measuring the Implementation of Behavioral Intervention Technologies: Recharacterization of Established Outcomes. J Med Internet Res. Journal of Medical Internet Research; 2019;21(1):e11752. PMID: 30681966

77. Klein JP, Knaevelsrud C, Bohus M, Ebert DD, Gerlinger G, Günther K, et al. Finding the needles in the haystack: Quality criteria for the use of internet-based self-management interventions in the prevention and treatment of mental disorders. Fortschritte der Neurol Psychiatr. © Georg Thieme Verlag KG; 2019;87(3):187–191.

Supplementary Files

Figures

CONSORT (or other) checklists

Multimedia Appendixes