Man vs. machine: A meta-analysis on the added value of human support in text-based internet treatments ("e-therapy") for mental disorders

Koelen, J.A.; Vonk, Anne; Klein, A.; de Koning, L.; Vonk, P.; de Vet, S.; Wiers, R.

DOI
10.1016/j.cpr.2022.102179

Publication date
2022

Document Version
Final published version

Published in
Clinical Psychology Review

Clinical Psychology Review 96 (2022) 102179

Contents lists available at ScienceDirect

Clinical Psychology Review

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

Review

Man vs. machine: A meta-analysis on the added value of human support in text-based internet treatments ("e-therapy") for mental disorders

J.A. Koelen [a,*], A. Vonk [a], A. Klein [a], L. de Koning [a], P. Vonk [b], S. de Vet [a], R. Wiers [a,c,d]

[a] Developmental Psychology, Department of Psychology, University of Amsterdam, the Netherlands
[b] Department of Research, Development and Prevention, Student Health Service, University of Amsterdam, the Netherlands
[c] Center for Urban Mental Health, University of Amsterdam, the Netherlands
[d] Addiction Development and Psychopathology (ADAPT)-Lab, Department of Psychology, University of Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Text-based internet treatment
Technological guidance
Human guidance
Optional support
Therapist qualification
Mental disorders

ABSTRACT

Guided internet-based treatment is more efficacious than completely unguided or self-guided internet-based treatment, yet within the spectrum of guidance, little is known about the added value of human support compared to more basic forms of guidance. The primary aims of this meta-analysis were: (1) to examine whether human guidance was more efficacious than technological guidance in text-based internet treatments ("e-therapy") for mental disorders, and (2) whether more intensive human guidance outperformed basic forms of human guidance. PsycINFO, PubMed and Web of Science were systematically searched for randomized controlled trials that directly compared various types and degrees of online guidance. Thirty-one studies, totaling 6215 individuals, met inclusion criteria. Results showed that human guidance was slightly more efficacious than technological guidance, both in terms of symptom reduction ($g = 0.11$; $p < .01$) and adherence ($0.26 < g < 0.29$; p's < 0.01). On the spectrum of human support, results were slightly more favorable for regular guidance compared to optional guidance, but only in terms of adherence (OR = 1.89, $g = 0.35$; $p < .05$). Higher qualification of online counselors was not associated with efficacy. These findings extend and refine previous reports on guided and unguided online treatments.

1. Introduction

An increasing number of studies have shown that internet-based interventions are efficacious for a variety of mental disorders, such as anxiety disorders, depression, and problematic alcohol use (Andersson & Cuijpers, 2009; Domhardt, Geßlein, von Rezori, & Baumeister, 2019; Hadjistavropoulos, Mehta, Wilhelms, Keough, & Sundström, 2020). With internet-based treatment, we refer to a specific form of online intervention that makes use of texts, images, and videos to provide the client with therapeutic material in an interactive way, often in the shape of a fixed number of sequential modules, consisting of psycho-education, in-session exercises, and homework assignments. This predominantly text-based intervention is referred to as "e-therapy" throughout this manuscript. The effects of e-therapy are similar to those found for face-to-face therapy, at least when restricted to cognitive-behavior therapy (CBT; Andersson & Titov, 2014; Carlbring, Andersson, Cuijpers, Riper, & Hedman-Lagerlöf, 2018; Cuijpers, Donker, Van Straten, Li, & Andersson, 2010). Moreover, some follow-up studies have indicated that the effects of e-therapy are maintained for as long as five years after treatment (Hedman et al., 2011). There is burgeoning evidence that e-therapy, despite high initial costs, could be a cost-effective treatment, both as stand-alone treatment, or as an initial treatment option within a stepped-care model (Salivar, Rothman, Roddy, & Doss, 2020; Weisel, Zarski, Berger, Krieger, Schaub, et al., 2019).

E-therapy offers many other benefits for people with mental health issues, as they have the potential to overcome barriers to regular mental health services (Andersson, 2015; Lovell & Richards, 2000). For example, it may provide people living in remote or underprivileged areas with the opportunity to gain access to mental health care. Across the globe, nearly 10% of the world population have to travel for over an hour to reach the help they might need (Weiss et al., 2020). In addition to providing more flexibility and autonomy, some people may prefer to receive treatment in the privacy of their homes, likely related to stigma surrounding mental health problems (Andersson, Titov, Dear, Rozental, & Carlbring, 2019). Moreover, e-therapy, when implemented on a large scale, offers great potential for the prevention of mental disorders

* Corresponding author at: Nieuwe Achtergracht 129B, 1018 WT Amsterdam, the Netherlands.
E-mail address: jurkoel@gmail.com (J.A. Koelen).

https://doi.org/10.1016/j.cpr.2022.102179
Received 3 September 2021; Received in revised form 28 April 2022; Accepted 4 June 2022
Available online 9 June 2022

(Deady et al., 2017). Finally, during the recent COVID-19 pandemic, online treatment (including remote face-to-face therapy) has offered many people access to therapy without running the risk of contracting the virus. In summary, e-therapy increases access to mental health care across the globe and offers some important advantages that make this a promising avenue for future mental health care.

1.1. Guidance within e-therapy

In research contexts, many of the e-therapies studied have been "unguided" (sometimes called "self-guided"), where the client progresses through the treatment without any assistance. Increasingly, research is becoming available that better mirrors clinical practice, in which a therapist or counselor guides the client through the treatment (Andersson, 2015). Online guidance can be delivered in an asynchronous way, for example with regular individualized or semi-standardized email feedback, or in a synchronous way, through brief phone or chat sessions (Riper & Cuijpers, 2016). Videoconferencing is becoming increasingly popular, but is not the focus of this review, because this could be considered face-to-face treatment, and is usually sharply demarcated from modular internet-based treatments, that are the focus of this review (e.g., Berryhill, Culmer, Williams, Halli-Tierney, Betancourt, et al., 2019). The focus of this review is on guided e-therapy.

Guidance offers the possibility to tailor the therapy to the individual's needs, and to intervene better in case of non-adherence, crisis, or after a sudden increase in symptoms (Andersson & Titov, 2014). Most meta-analyses suggested that guidance renders e-therapy more efficacious (Andersson & Cuijpers, 2009; Johansson & Andersson, 2012; Richards & Richardson, 2012; Spek et al., 2007; Van 't Hof, Cuijpers, & Stein, 2009). However, these meta-analyses did not compare guided and unguided e-therapy directly with each other. Instead, they compared one group of studies that contrasted guided e-therapy with treatment as usual (TAU), with another group of studies that contrasted unguided e-therapy with TAU. The finding that the effect sizes for guided e-therapy was larger than those for unguided led to the conclusion that guided e-therapy was more efficacious. However, given that guided and unguided e-therapies were not offered within the same setting, this conclusion should be confirmed in analyses with direct comparisons.

Three more recent meta-analyses with head-to-head comparisons tended to confirm the superiority of guided e-therapy for a range of mental disorders (Baumeister, Reichler, Munzinger, & Lin, 2014; Domhardt et al., 2019; Karyotaki et al., 2021). For example, in a meta-analysis of 8 studies into a variety of mental disorders (mostly social phobia and depression), Baumeister et al. (2014) reported a standardized mean difference of $d = 0.27$ in favor of the guided e-therapy. Domhardt et al. (2019), examining the efficacy of e-therapy for anxiety disorders, found a similar effect size difference of $d = 0.39$ in favor of guided e-therapy. However, this result was based on only four included studies. Karyotaki et al. (2021), in an individual patient data network meta-analysis including 39 studies on e-therapy for depression, reported moderate differences ($d = 0.6$) between guided and unguided treatments, in favor of the guided conditions. These findings were stronger for patients with higher depression scores (they benefited more from the guided treatment). However, the differences disappeared at 6- or 12 months following randomization, although it should be noted that the latter finding was based on a subgroup of 8 studies only (Karyotaki et al., 2021). Findings were based on online CBT in patients with depression only, and treatment duration (i.e., dosage of guidance) was not taken into account.

A significant concern is that despite superficial agreement between these meta-analyses, basic forms of support are sometimes included in the so-called "unguided" treatment conditions (e.g., Karyotaki et al., 2021). Newer unguided e-therapies differ in the sense that participants often do receive automated messages intended to increase adherence and to reinforce their progression through treatment, which was much less the case in older forms of e-therapy (Dear, Staples, Terides, Fogliati, Sheehan, et al., 2016). This could be considered "technological support", yet was not taken into account in some of the older meta-analyses and reviews. Riper et al. (2018), for example, found that human-supported e-therapies were more efficacious to reduce problem drinking than "fully automated" ones, yet it was not made explicit what the automated interventions entailed. To elucidate these issues, in this review we will make a clear distinction between fully unguided and technologically guided treatments. As noted, the comparison between guided and fully unguided e-therapy was the focus of other reviews; our focus lies on the full spectrum of guidance, consisting of technological guidance at the one end, and varieties of human guidance at the other end of the spectrum.

1.2. Varieties of human guidance

Within the spectrum of human guidance, a further distinction can be made between intensive and more basic human guidance (Domhardt et al., 2019; Newman et al., 2011; Richards & Richardson, 2012). In the aforementioned meta-analysis of anxiety disorders, Domhardt et al. (2019) differentiated between "guided" and "mostly unguided" interventions. Treatment was considered mostly unguided when "technical support" was offered at the request of the patient. Please note that technical support (i.e., a human being helping to solve a technical issue) should be differentiated from "technological" support mentioned earlier, which is non-human by definition. However, to complicate matters further, in some studies the "technical" support refers to scheduled, motivational support, and encouragement from psychologists (e.g., Dirkse et al., 2020; Johnston et al., 2011); a component considered a "common factor" of effective therapies (Cuijpers, Reijnders, & Huibers, 2019; Wampold & Imel, 2015). In other studies, "technical support" is provided by non-psychologists on a weekly basis to encourage and motivate participants (e.g., Titov, Andrews, Davies, Mcintyre, Robinson, et al., 2010). Richards and Richardson (2012), in their meta-analysis, differentiated between studies that offered therapist support with those that offered "administrative" support, which appears similar to some definitions of "technical support". From these examples, it becomes clear the type/degree of human support, and the qualification of the person supporting the treatment are sometimes conflated. Researchers do not appear to agree on what is meant by "technical" or "administrative" support. In our opinion, more clear definitions of degrees of guidance in e-therapy, as well as the distinction between intensity of guidance and qualification of counselors, are needed in order to analyze and understand their effect.

Therefore, in this meta-analysis, we compared varying degrees of human support that were restricted to clinical guidance, i.e., support aimed at the content of the program and not at its usage, using a clearly defined taxonomy. This spectrum of human guidance includes three levels: (1) Minimal human guidance, excluding mere assistance for technical problems. Minimal guidance refers to support on demand, i.e., optional support is provided only when the patient asks for it; (2) Regular (scheduled) guidance in the form of e-mail feedback (asynchronous) to assignments or questions, or brief support via telephone or chat (synchronous). Regular guidance followed the established regime of planned weekly support, and (3) Intensive guidance, i.e., human support that is offered more frequently (a fixed higher frequency of contact, i.e., 2 or 3 times a week), or more quickly (e.g., within 24 h), than regular support. We realize that optional support (level 2) could in practice be more intensive than level 3 type of support (i.e., when there is high demand for it). Therefore, studies offering optional vs. regular support were also analyzed separately, to control for the potential confound of intensity. Studies with a focus on levels of counselor qualification were compared separately from the matter of intensity, to allow for a comparison of high-qualified guidance from low-qualified guidance (Baumeister et al., 2014).

1.3. Definition of technological guidance

All types of guidance share the common aim of guiding the patient through online treatment modules, and increasing adherence (Andersson, 2015; Riper and Cuijpers, 2016). Technological guidance in our conceptualization consisted of automated reminders and feedback or encouragement. "Reminders" imply messages to inform participants about new material available, additional resources, or the aim to instigate planning exercises. These messages were usually sent at fixed intervals, or when participants were unresponsive. "Automated feedback" or encouragement/reinforcement refers to automatic standardized (template-based) messages that the participant receives upon session completion, usually to congratulate with completion of the session and thus reinforce progress, and/or to provide a summary of the contents.

1.4. Aims of the present study

In sum, e-therapy studies use a wide variety of definitions to refer to the type and nature of guidance being offered. Yet, most reviews and meta-analyses used a dichotomization by comparing the coarse categories of 'guided' and 'unguided' interventions, which fails to consider the wide spectrum of guidance and its variations (Farrand & Woodford, 2013). Moreover, previous meta-analyses have rarely made head-to-head comparisons. These two issues render it difficult to draw definite conclusions with respect to which type of support is optimal.

The primary aims of this study were to firstly clarify whether human guidance would increase efficacy compared with technological guidance only, and secondly whether more intensive human guidance would increase efficacy of e-therapy compared with more basic forms of human guidance. To address these aims, we created two separate sets of comparisons. Our first set of comparisons was between studies that directly compared technological and human guidance. Our second set of comparisons concentrated solely on varieties within the spectrum of human guidance, in which we differentiated between three levels of human guidance, as introduced above. This approach differs from previous meta-analyses and reviews in three ways: (1) We did not include interventions that were completely self-guided, as well as studies with technical support only; (2) Compared to 'regular guidance' (weekly human support), we included both less and more intensive forms of human guidance; (3) We included only studies directly comparing varieties of guidance.

Furthermore, due to the confusion between qualification of the online counselor and so-called "technical" or "administrative support" (e.g., Dirkse, Hadjistavropoulos, Alberts, Karin, Schneider, et al., 2020; Richards & Richardson, 2012; Titov, Andrews, Schwencke, Solley, & Robinson, 2009), we compared studies that examined the impact of therapist qualification on outcome separately, thus updating previous reviews (e.g., Baumeister et al., 2014). Finally, a number of moderators were examined, such as offering a pretreatment interview or actively reminding participants of their assignments.

2. Methods

2.1. Literature search

An initial, systematic multi-phase search was conducted May 2020 in three databases (PsycINFO, PubMed and Web of Science) to obtain studies that reported on the impact of therapist guidance in e-therapy (see Supplement 1, Appendix A for the search strategy). This search was updated December 2021. Our meta-analysis focused on e-therapy, and not on combinations of face-to-face and e-therapy (blended therapy).

Publication year of published articles was not constrained. Within the domain of randomized controlled trials (in English), we used the following search terms (see also Appendix A in Supplement 1): web-based, online, internet*, digital* or computer* together with cognitive behav* or therap* or treatment, and assistance, support or guidance in conjunction with various qualifications of guidance. To detect recently completed trials, registered trials in the U.S. National Library of Medicine (https://www.clinicaltrials.gov) were searched. In case (published) results were to be expected, researchers were contacted to obtain potential results to be included in this meta-analysis. This yielded no additional studies. Authors were also contacted in case of incomplete or missing data. This meta-analysis was pre-registered (PROSPERO 2021 CRD42021243964).

2.2. Inclusion and exclusion criteria

Randomized controlled trials were included if they fulfilled the following criteria:

(1) adult participants (18+);
(2) a mental disorder according to either relevant classification systems or a subthreshold disorder, using a validated cut-off (screener), or both. The disorder or the dimensional equivalent had to be enlisted in the official handbooks of mental disorders (DSM-IV or 5, ICD-10 or 11);
(3) the outcome of the intervention was assessed in terms of depression, anxiety, or both. Thus, sleep disorders, sexual disorders, and somatic symptom disorders were also included, as long as the focus was on the alleviation of depression and/or anxiety;
(4) publication in English;
(5) examination of variations of therapist guidance in internet treatment with at least two guided interventions with different intensities (e.g., regular or optional, high or low frequency) of guidance. Studies comparing different levels of therapist qualification were also included;
(6) trials had to report (a) symptom (depression/anxiety) severity levels at posttreatment or (b) adherence to the program as outcomes (or both). Adherence was operationalized following Donkin et al. (2011) as the percentage of participants that completed the whole treatment, and as the mean number of sessions completed.

Studies were excluded if they:

(1) contained no e-therapy as defined here (e.g., attentional bias modification training, psychoeducation only, cognitive or physical remediation therapy);
(2) combined e-therapy with face-to-face therapy (blended therapy), either simultaneously or sequentially, or only face-to-face treatment, or face-to-face treatment as a control group;
(3) examined only "self-guided" treatments with no form of guidance. Note that all studies that claimed to examine "self-guided treatments" were scrutinized for the actual absence of guidance in any shape or form (technological support), as newer types of internet therapies often provide automated support in "self-help" interventions;
(4) static webpages offering psychoeducation only;
(5) comparison of two types of treatment with the same level of guidance;
(6) inclusion of fully automated programs with virtual therapists or chatbots, with unlimited access to the program;
(7) test of therapeutic effects of programs using virtual or augmented reality or games;
(8) inclusion of supportive communication or supportive therapy as a control ("attention control") without any guidance of modules (e.g., e-mails alone);
(9) inclusion of non-moderated internet forums as main 'intervention' platform. Forums were allowed when offered in addition to modules, and if moderated (and not just *monitored*) by a clinical psychologist.

2.3. Selection of studies

The studies were selected in two phases: (1) screening of title and abstract and (2) inspection of full text. To complement the electronic search, reference lists of recent meta-analyses and reviews on this topic (i.e., Baumeister et al., 2014; Domhardt et al., 2019; Karyotaki et al., 2021) were screened for relevant articles during the first phase. In addition, reference lists of the screened full-text papers were inspected when deemed relevant.

Two researchers (JK and AV) independently assessed the inclusion and exclusion criteria after an initial calibration. Both authors screened all retrieved search hits. A conservative approach was taken, so when the title and abstract did not provide enough information, the article was inspected full text. In the first phase, the agreement between the two raters was 92% (Cohen's kappa = 0.62), which is considered substantial agreement. For the second phase, the agreement was 74% (Cohen's kappa = 0.48), which is considered moderate agreement. The main reason the agreement dropped in the second phase was a lack of clarity about what constituted guidance, and which control groups were allowed (e.g., supportive therapy without any modules). These issues were refined during consensus meetings and yielded the definitions introduced above. Disagreement was resolved by discussion until consensus was reached. There was no need to consult a third party to reach consensus.

2.4. Data extraction

For each study included, the same raters extracted the statistics necessary for effect size calculation (means, standard deviations, drop out or adherence rate, sample sizes) for the relevant treatment conditions and the relevant outcome data. The primary outcomes differed per study and were usually determined by the main disorders under treatment. Psychological symptoms were chosen as outcome, also for studies in the realm of medical psychology. Effect sizes were calculated for psychological symptoms and for adherence to treatment (see definition above). Post-treatment scores were obtained where available within three months of treatment completion. Because follow-up outcome periods are likely to vary across studies, and because we were interested in the immediate impact of guidance, we focused only on post-treatment outcomes. Self-report measures were included as most studies use self-report instruments only.

Finally, study characteristics were extracted (or calculated), that could be used as moderators, including primary diagnosis or complaint, setting (community/website, primary care, clinic, or hospital), type of treatment (i.e., CBT or not), number of sessions or modules, and therapist qualification (level of training, and/or role).

2.5. Assessment of study quality

To determine the methodological quality of included studies, they were rated with the RCT Psychotherapy Quality Rating Scale (RCT-PQRS; Kocsis et al., 2010). After registration, but before data-extraction, we decided to use this instrument instead of the Cochrane risk of bias tool (Higgins & Green, 2011), because the PQRS is better tailored to the particularities of (psycho)therapy (e.g., that clinician and patient are not blind to the treatment provided). The RCT-PQRS was specifically developed for RCTs in psychotherapy research and contains 25 items covering six domains: (a) description of patients; (b) definition and delivery of treatment; (c) outcome measures; (d) data analysis; (e) treatment assignment; and (f) overall quality. The last 'omnibus' item is scored on a 7-point scale; other items on a 3-point scale (0–2), yielding a range of 1–55, with scores $\leq$9 representing abominable quality, scores 10–14 very poor quality, 15–24 poor quality, 25–33 adequate quality, 34–42 good, 43–50 very good, and $\geq$ 51 excellent quality.

One independent judge (a Master psychology student) trained by the first author coded all studies. To establish interrater reliability, the first author rated a random sample of 9 studies. Intraclass correlations (ICC) coefficients were calculated using SPSS Statistics for MacIntosh, version 24 (IBMCorp., 2018), based on a mean rating (k = *2)*, absolute-agreement, 2-way mixed-effects model. The intraclass correlation of single measures was 0.84, which indicates a good reliability (Koo & Li, 2016).

2.6. Data analysis

2.6.1. Computation of effect sizes

2.6.1.1. Calculations of between-group contrasts. The post-treatment scores for the two conditions that were being compared were contrasted and divided by their pooled standard deviation [$M_1 - M_2$/ sd_{pooled}]. First, we provided a global estimate for between-group contrasts across all studies, generalizing across types of guidance. Second, subgroups addressing frequency or speed of feedback were analyzed separately from those that compare regular vs. optional guidance. When a study included multiple outcomes, the means of z-transformed variables were used to calculate an average effect size per study. This approach yields a conservative estimate, because the correlation for the separate outcomes per study is assumed to be 1 (while in reality it will be lower) (Borenstein, Hedges, Higgins, & Rothstein, 2009). Some studies yielded more than one effect size, because they contained more than two treatment groups. In this case, we considered these pairwise comparisons separately. To avoid "double counts" in the shared intervention group (that served as the comparison), the shared group N was split in half (Higgins & Green, 2011).

2.6.1.2. Computation of pooled effect sizes across studies. Meta-analyses were performed using Comprehensive Meta-Analysis (Borenstein, Hedges, Higgins, & Rothstein, 2005). Standardized mean differences with 95% confidence intervals (95% CI) were computed for all continuous outcomes. Hedges' g was used because this corrects for small sample sizes (Hedges & Olkin, 1985). Effect sizes of 0.20, 0.50 and 0.80 are considered small, medium, and large (Cohen, 1988). For dichotomous variables, odds ratios (OR) with 95%CI were computed. Positive effect sizes imply that higher levels of guidance yielded higher effect sizes.

A random-effects model was used to compute weighted mean effect sizes, because we expected true population effect sizes to vary across studies due to differences in sample, methodology and treatment. The random-effects model results in more conservative results and broader 95%CI than the fixed-effects model. With this procedure, effect sizes are weighted by their inverse variance, thus giving more weight to larger studies (with smaller sampling error) and increasing the reliability of the effect estimates. To examine the robustness of the global effects, we employed the 'one study removed' method. Furthermore, effects were recalculated without outliers. A study was judged an outlier when the confidence interval of the study did not overlap with the pooled effect size (Harrer, Cuijpers, Furukawa, & Ebert, 2021). Finally, effects were recalculated for studies with data for the full randomized sample (intention-to-treat sample, or ITT). ITT samples usually give more conservative estimates of relative treatment effects, especially when dropout is high, as is often the case in internet treatments.

2.6.1.3. Heterogeneity. Heterogeneity of effect sizes within and between subsamples were calculated using the Q and the I^2 statistic (Higgins & Thompson, 2002). Significant p-values for the Q test indicate the presence of heterogeneity. I^2 represents the percentage of total variance in effect estimates that is due to systematic heterogeneity between studies rather than due to chance or sampling error. Low percentages indicate low heterogeneity and percentages above 75% substantial heterogeneity.

2.6.1.4. Moderator analysis. For the purpose of moderator analyses,

studies were divided into subgroups. For each subgroup the pooled mean effect size was calculated, and differences in effect sizes between the subgroups (with a minimum of four studies) were examined for statistical significance using the *Q* statistic. For the comparison of subgroups, the mixed-effects model was used. This model uses the random-effects models to estimate the effect size for each subgroup, while the fixed-effects model is used to test the difference between the subgroups (Borenstein et al., 2009).

The following moderators were examined: (1) studies providing a pre-treatment interview were analyzed separately and compared to studies without, because, based on previous findings (Boettcher, Berger, & Renneberg, 2012), we hypothesized that a pre-treatment screening or motivational session would decrease the between-group effect sizes, and may outweigh the effect of guidance during treatment (Johansson & Andersson, 2012); (2) we compared studies that offered reminders in both treatment conditions, to those that did not, as this likely decreases between-group differences; (3) studies were analyzed separately for those that treated anxiety disorders, and compared to those that did not, as internet treatments for anxiety disorders show inconsistent findings, and the desired level of guidance is unclear (Farrand & Woodford, 2013; Spek et al., 2007); (4) studies offering CBT were analyzed separately and compared to those with another therapeutic orientation; (5) studies were grouped according to their mode of delivery. We distinguished between 'synchronous' communication mode (chat, telephone), 'asynchronous' communication mode (email), and mixed, in line with other meta-analyses (Baumeister et al., 2014).

2.6.1.5. Publication bias. We tested potential publication bias by means of the iterative non-parametric trim and fill procedure as implemented in CMA. This procedure controls for the association between individual effect sizes and their sample sizes (i.e., sampling error) by inspecting funnel plots. Publication bias is assumed to be present when the effect sizes of small studies - with larger sampling variation than large studies - are represented asymmetrically within and around the funnel (Sterne & Egger, 2001). The Duval and Tweedie procedure (Duval & Tweedie, 2000) provides a correction of the effect size after publication bias has been taken into account by trimming away studies suggesting asymmetry. We used the random-effects model. In addition, we used Egger's regression intercept (Egger, Smith, Schneider, & Minder, 1997) and Begg and Mazumdar's (1994) rank correlation test.

3. Results

The electronic database search yielded 1629 hits, and 24 additional records were identified through other sources (online registers, cross-references, etc.). After removal of duplicates, 1272 articles remained.

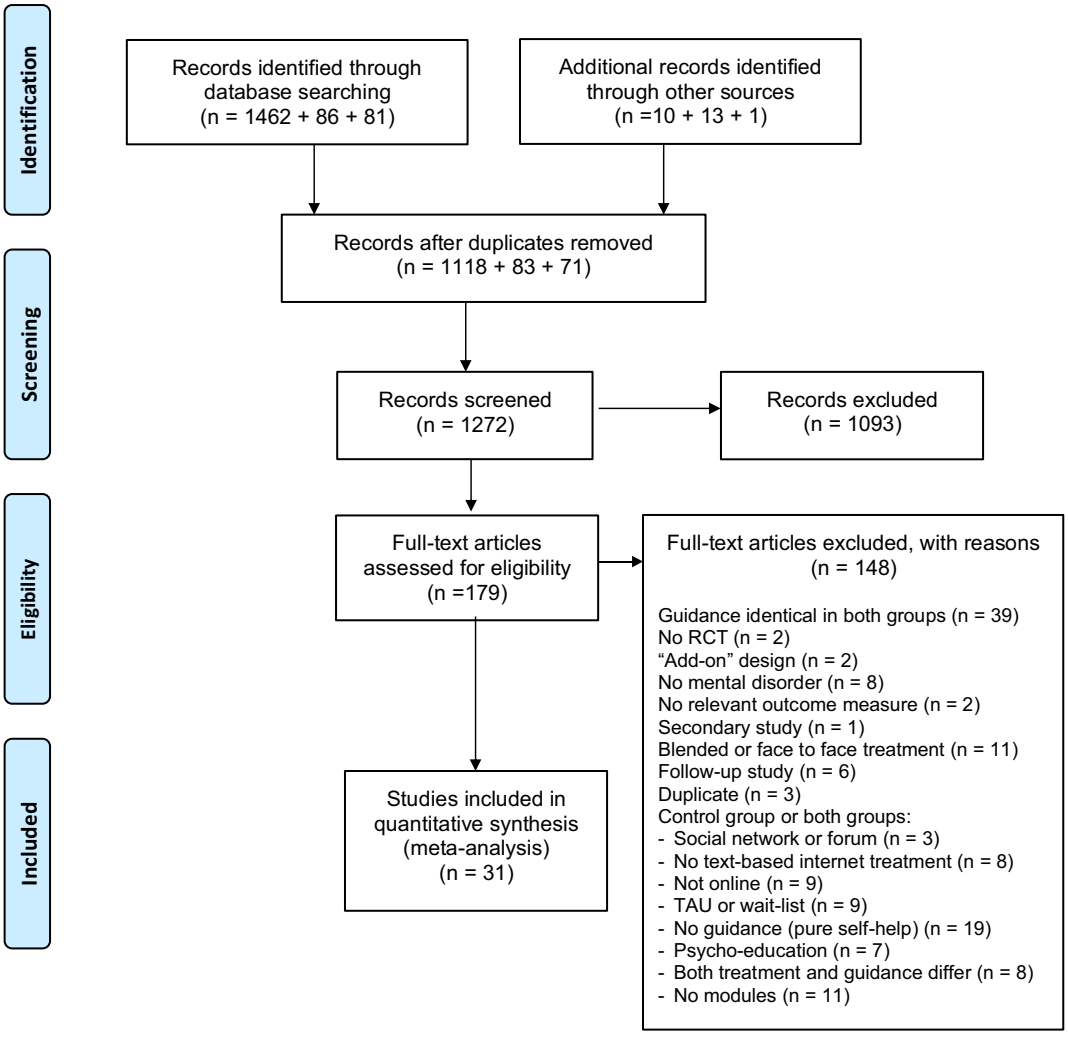

Fig. 1. Flow diagram of screening process.
Note. An original search (May,2020)and two subsequent searches were conducted (March and December 2021). This is depicted in the diagram by first mentioning the results from the original search, and then (with+) the results from the second and third searches. All searches were identical in that they were conducted with the same search strings, in the same databases, and the by the same author (JK).

Screening of title and abstract resulted in the exclusion of 1093 articles. The remaining 179 articles were scrutinized full text. During this phase, a total of 148 studies were excluded (see Fig. 1 for reasons), leaving 31 studies to be included in the meta-analysis, totaling 6215 participants. The study selection process is detailed in the PRISMA flow chart (Fig. 1).

3.1. Characteristics of the included studies

Appendix B (Supplement 1) provides a summary of the included studies and their main characteristics. Two publications turned out to stem from one original study, albeit with different outcome measures (Brabyn et al., 2016; Gilbody et al., 2017). Thus, 30 unique studies were included, with 32 potentially relevant comparisons (two studies had three relevant treatment conditions). Eleven studies compared various degrees of human guidance, fourteen studies compared human guidance with technological guidance. Two studies could not be allocated to our pre-defined comparisons. Schulz et al. (2016) compared group (forum) versus individual treatment. Another study (Sundström et al., 2016) compared different modalities of guidance (choice of fixed chat or email) versus therapist contact through email. Seven studies compared different levels of therapist qualification. From another study (Pihlaja et al., 2020) symptom reduction could not be obtained, yet data from adherence measures was available.

All but five studies (83.3%) treated anxiety and/or depression. Thirteen studies (43.3%) addressed panic disorder, flying phobia, social anxiety disorder, generalized anxiety disorder or severe rumination (Berger et al., 2011; Campos et al., 2019; Cook et al., 2019; Dear et al., 2015, 2016; Fogliati et al., 2016; Ivanov et al., 2016; Johnston et al., 2011; Klein et al., 2009; Oromendia et al., 2016; Robinson et a., 2010; Schulz et al., 2016; Titov et al., 2009). Eight studies (26.7%) treated depressive symptoms (Farrer et al., 2011; Gilbody et al., 2017; Mohr et al., 2013; Montero-Marin et al., 2016; Pihlaja et al., 2020; Titov et al., 2010; Westerhof et al., 2019; Zagorscak et al., 2018), and four studies (13.3%) targeted a combination of anxiety and depression (Hadjistavropoulos et al., 2017; Hadjistavropoulos, Peynenburg, Nugent, et al., 2020; Hadjistavropoulos, Peynenburg, Thiessen, et al., 2020; Kleiboer et al., 2015). Three other studies focused on other disorders: insomnia (Lancee et al., 2013), severe symptoms of eating disorders (Aardoom et al., 2016), and problematic alcohol use (Sundström et al., 2016). The remaining two studies were conducted in medical settings, and focused on psychological symptoms in cancer survivors (Dirkse et al., 2020), and haemodialysis patients (Hudson et al., 2017). The latter two studies used a threshold to screen for depression and anxiety as part of the inclusion process.

All but 4 studies (86.7%) offered CBT; the median number of modules was 6 (range: 5–18). One study offered Acceptance and Commitment Therapy (Ivanova et al., 2016), one problem-solving therapy (Kleiboer et al., 2015), and one study provided life-review therapy (Westerhof et al., 2019). For one study, the type of therapy was unclear (Aardoom et al., 2016). More than half of the studies ($k = 17$; 56.7%) were conducted in the community: they recruited through websites or newspapers. Other studies made use of a combination of the community, websites, and mental health care settings ($k = 5$; 16.7%). Three studies (10.0%) were conducted in hospital settings (Hudson et al., 2017; Mohr et al., 2013; Pihlaja et al., 2020), and one made use of specialized mental health care facilities (Farrer et al., 2011). The remainder of the studies (13.3%) were conducted in primary care (Gilbody et al., 2017; Montero-Marín et al., 2016), university (Cook et al., 2019), and one made use of the information from the archives of insurance companies (Zagorscak et al., 2018). Studies were conducted in Australia, Canada, Finland, Germany, the Netherlands, the United Kingdom, Spain, Sweden, Switzerland, and the United States.

Most studies used an ITT analysis format and imputed data from missing cases. However, upon closer inspection, several studies only analyzed the data of participants who started treatment. For the purpose of simplicity, this type of analysis is regarded as "modified ITT" analysis. For studies that reported data from completers only, authors were e-mailed. Several authors were able to provide us with the estimated averages for the entire sample. For the adherence rates, all 28 studies with relevant comparisons had adherence data on at least one of the two outcomes. For approximately one quarter of studies with missing data on one of the outcomes, data could still be obtained for the other outcome.

3.2. Methodological quality of the included studies

Results from the quality ratings are reported in Appendix C (Supplement 1). Half of the studies were rated as good, 6 studies (20.0%) were rated as very good, and 9 studies (30.0%) were rated as adequate in terms of methodological quality.

3.3. Comparison 1: human vs. technological guidance

Before conducting these analyses, we checked for the degree of sophistication in the technological support conditions, because in theory, automated support could be very sophisticated and tailor-made and potentially more frequently available for patients. Upon closer inspection, we found one study with a high degree of sophistication, using a feedback algorithm based on 4 dimensions of symptom severity (Aardoom et al., 2016). The other studies used fixed templates for their feedback. In light of this, we also analyzed the subset of studies related to the comparison of human versus technological guidance separately without the advanced feedback study.

Fourteen studies were available regarding this comparison for the outcome of symptoms. The pooled effect size was $g = 0.11$ (95% CI: 0.03, 0.19; $p < .01$) indicating that human guidance was slightly, yet significantly, more efficacious than automated guidance (see Table 1). Heterogeneity was absent and non-significant ($I^2 = 0\%$; $Q(13) = 7.48$; $p = .88$). Using the one-study removed method yielded effect sizes in the range of $g = 0.09$–0.12. Analyses including only modified ITT-data ($k = 6$) yielded similar outcomes (Table 1). *Re*-analyzing the data without the study with advanced technological guidance yielded similar findings ($g = 0.12$; 95% CI: 0.04, 0.20; $p < .01$). In three studies (Hudson et al., 2017; Ivanova et al., 2016; Montero-Marin et al., 2016), the frequency of human support deviated from the standard frequency of once per week. We also analyzed the subset without these studies, which yielded similar results: $g = 0.11$; 95% CI: 0.03, 0.19; $p < .01$).

For the outcome of mean number of sessions completed, 9 studies provided the required data (Table 2). This analysis yielded a pooled effect size of $g = 0.26$ (95% CI: 0.13, 0.40; $p < .01$), indicating that individuals receiving human support completed more sessions on average. Heterogeneity was moderate and significant ($I^2 = 52.1\%$; $Q(8) = 16.71$; $p < .05$). There were no outliers. Using the one-study removed method, we observed effect sizes between 0.18 and 0.29 (all p's < 0.01). Excluding the studies with divergent frequencies of human support (Ivanova et al., 2016; Montero-Marin et al., 2016), yielded a similar outcome: $g = 0.25$; 95% CI: 0.09, 0.40; $p < .01$).

In terms of adherence rates, those receiving human guidance were more likely to complete treatment (OR = 1.69; 95%CI: 1.30, 2.19; $p < .01$). Heterogeneity was moderate but non-significant (Table 3). Using the one-study removed method, we found odds ratios between 1.52 and 1.77 (all p's < 0.01). Removing one outlier (Lancee et al., 2013) yielded a somewhat lower but still significant odds ratio of 1.52 (95%CI: 1.27, 1.83; $p < .01$). The outcome was similar without studies with varying frequencies of human support (Hudson et al., 2017: Ivanova et al., 2016; Montero-Marin et al., 2016): $OR = 1.72$; 95% CI: 1.26, 2.35; $p < .01$.

We performed moderator analyses only for outcomes with sufficient studies (symptom outcomes). For the moderators "pre-treatment interview", "anxiety disorder", and "mode of delivery", non-significant differences between the designated subgroups were found. For the moderators "reminders in both groups", and "CBT vs. other treatment" not enough studies were available in each subgroup to allow for meaningful comparisons. Most studies were CBT-based and most offered

Table 1
Results for human guidance vs. technological guidance – symptoms.

		95% Confidence interval				Heterogeneity			
Study name	Hedges'*g*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Aardoom2016 (1)	0.00	−0.30	0.29	−0.02	0.99				
Campos2019	0.12	−0.56	0.80	0.35	0.73				
Cook2019	−0.12	−0.48	0.24	−0.66	0.51				
Dear2015	0.02	−0.29	0.34	0.15	0.88				
Dear2016	0.08	−0.30	0.46	0.42	0.67				
Dirkse2020	0.15	−0.28	0.58	0.67	0.50				
Fogliati2016	0.02	−0.50	0.54	0.08	0.93				
Hudson2017	0.07	−0.79	0.93	0.16	0.87				
Ivanova2016	0.29	−0.10	0.68	1.45	0.15				
Kleiboer2015(2)	0.12	−0.15	0.38	0.84	0.40				
Lancee2013	0.32	0.08	0.57	2.62	0.01				
Mohr2013	−0.02	−0.48	0.45	−0.06	0.95				
Montero-Marin2016	−0.05	−0.40	0.31	−0.26	0.80				
Zagorscak2018	0.13	0.01	0.24	2.07	0.04				
Weighed mean *g* (random effects)	**0.11**	**0.03**	**0.19**	**2.79**	**<0.01**	**7.48**	**13**	**0.88**	**0%**
ITT only (k = 6)	0.14	0.05	0.23	3.12	<0.01				

Table 2
Results for human guidance vs. technological guidance – session average completed.

		95% Confidence interval				Heterogeneity			
Study name	Hedges'*g*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Campos2019	0.03	−0.54	0.59	0.09	0.93				
Dear2015	0.26	0.04	0.47	2.36	0.02				
Dear2016	0.08	−0.18	0.34	0.59	0.56				
Dirkse2020	0.49	0.07	0.92	2.27	0.02				
Fogliati2016	0.07	−0.26	0.39	0.40	0.70				
Ivanova2016	0.48	0.08	0.88	2.36	0.02				
Lancee2013	0.61	0.36	0.86	4.85	<0.01				
Montero-Marin2016	0.22	−0.14	0.58	1.21	0.23				
Zagorscak2018	0.15	0.03	0.27	2.46	0.01				
Weighed mean *g* (random effects)	**0.26**	**0.13**	**0.40**	**3.79**	**<0.01**	**16.71**	**8**	**0.03**	**52%**

Table 3
Results for human guidance vs. technological guidance – rate of patients completing all sessions.

		95% Confidence interval				Heterogeneity			
Study name	*OR*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Campos2019	0.81	0.17	3.78	−0.27	0.79				
Cook2019	1.82	0.92	3.62	1.72	0.09				
Dear2016	1.20	0.67	2.14	0.61	0.54				
Dear2015	1.42	0.91	2.23	1.54	0.12				
Dirkse2020	3.12	0.74	13.20	1.54	0.12				
Fogliati2016	1.05	0.52	2.11	0.13	0.90				
Hudson2017	0.39	0.02	7.64	−0.62	0.54				
Ivanova2016	1.60	0.70	3.66	1.12	0.26				
Kleiboer2015(2)	1.75	0.84	3.63	1.49	0.14				
Lancee2013	*4.40*	*2.60*	*7.44*	*5.52*	*< 0.01*				
Montero-Marin2016	1.62	0.77	3.44	1.26	0.21				
Zagorscak2018	1.66	1.24	2.23	3.40	0.01				
Weighed mean *OR*	**1.69**	**1.30**	**2.19**	**3.97**	**<0.01**	**18.95**	**11**	**0.06**	**42%**
ITT only (k = 5)	1.97	1.31	2.95	3.26	<0.01				
Outliers removed (k = 11)	1.52	1.27	1.83	4.49	<0.01				

Note. Study names in italics represent outliers.

reminders in both groups.

3.4. Comparison 2: degrees of human guidance

Next, we compared studies with varying degrees of human guidance on the three outcomes. In terms of symptoms, we calculated the pooled effect size for ten studies (Table 4). These studies consisted of those comparing regular vs. optional guidance (Berger et al., 2011; Farrer et al., 2011; Gilbody et al., 2017; Hadjistavropoulos et al., 2017; Kleiboer et al., 2015; Oromendia et al., 2016), and those that compared a (fixed) higher frequency (i.e., 2 or 3 times a week) or speed (i.e., within one business day) of contact with a standard frequency (Aardoom et al., 2016; Hadjistavropoulos et al., 2020a, 2020b; Klein et al., 2009). The pooled effect size was non-significant at $g = 0.05$ (95% CI: −0.04, 0.15; p = .27), indicating that higher levels of human guidance were not more efficacious than lower levels of human guidance in terms of symptom reduction. Heterogeneity was low and non-significant ($I^2 = 9.0\%$; Q (9) = 9.89; p = .36). Analyzing this subset again with one outlier ($k = 9$) removed or ITT data only ($k = 8$) yielded similar results.

In terms of mean number of sessions completed, 7 studies were

Table 4
Results for degrees of human guidance – symptoms.

		95% Confidence interval				Heterogeneity			
Study name	Hedges'*g*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Aardoom2016 (2)	−0.08	−0.37	0.21	−0.53	0.60				
Berger2011*	0.04	−0.49	0.57	0.15	0.88				
Farrer2011*	0.01	−0.42	0.44	0.05	0.96				
Gilbody2017*	0.24	0.00	0.49	1.95	0.05				
Hadjistavropoulos2017*	−0.09	−0.39	0.21	−0.60	0.55				
Hadjistavropoulos2020a	−0.01	−0.19	0.18	−0.07	0.95				
Hadjistavropoulos2020b	0.03	−0.13	0.19	0.39	0.69				
Kleiboer2015(1)*	0.07	−0.20	0.34	0.50	0.61				
Klein2009	0.08	−0.45	0.62	0.31	0.76				
*Oromendia2016**	*0.74*	*0.16*	*1.32*	*2.50*	*0.01*				
Weighed mean *g* (random effects)	**0.05**	**−0.04**	**0.15**	**1.10**	**0.27**	**9.89**	**9**	**0.36**	**9%**
ITT only (k = 8)	0.02	−0.07	0.12	0.44	0.66				
Outliers removed (k = 9)	0.04	−0.05	0.12	0.78	0.43				

Note. Studies with * compare regular with optional support.
Study names in italics represent outliers.

compared (Table 5). This yielded a pooled effect of $g = 0.30$ (95%CI: 0.07, 0.53; $p < .05$), which indicates that more intensive human guidance was more efficacious than lower levels of support in terms of adherence. Heterogeneity was high and significant ($I^2 = 75.1\%$; Q (6) = 24.13; $p < .001$). Using the one-study removed method, we observed effect sizes between $g = 0.20$ and $g = 0.37$ (all p's < 0.05). Repeating these analyses with ITT data only (k = 6) yielded similar results. Excluding one outlier (Oromendia et al., 2016) led to a somewhat lower effect size: $g = 0.20$ (95%CI: 0.02, 0.37; $p < .05$), and reduced the heterogeneity to a moderate level ($I^2 = 56.7\%$; Q (5) = 11.54; $p < .05$). The results for adherence rates (Table 6) were similar. Nine studies were analyzed together, which resulted in an OR = 1.57 (95%CI: 1.09, 2.25; $p < .01$). This suggests that individuals with higher levels of human guidance were more likely to complete treatment. Heterogeneity was moderate and significant ($I^2 = 54.2\%$; Q (8) = 17.46; $p < .05$). Using the one-study removed method, we observed odds ratios between OR = 1.28 and 1.75 (all p's < 0.05).

As noted, to control for potential confound of intensity, we re-analyzed degrees of human guidance without studies that compared intensive guidance with regular guidance, so that only those comparing optional vs. regular guidance remained (k = 4–6). This analysis (for symptoms) yielded a somewhat higher, yet non-significant Hedges' g of 0.12 (95%CI: −0.06, 0.30; p = .18) for standard guidance compared to optional guidance ($k = 6$). Heterogeneity was low and non-significant ($I^2 = 35.6\%$; Q (5) = 7.77; p = .170). For number of completed sessions, comparing regular guidance with optional guidance ($k = 4$) yielded a significant effect: $g = 0.46$ (95%CI: 0.02, 0.91; $p < .05$). Yet, again heterogeneity was high ($I^2 = 73.9\%$; Q (3) = 11.51; $p < .001$), so this estimate was not reliable. In terms of adherence rates ($k = 5$), we observed an OR = 1.89 (95%CI: 1.07, 3.34; $p < .05$), indicating that regular guidance was more efficacious than optional guidance.

For the remaining set of studies ($k = 4$) comparing higher frequency/speed with regular frequency/speed, the effect was non-significant: OR = 1.24 (95%CI: 0.83, 1.85; p = .29). Heterogeneity was low to moderate and non-significant ($I^2 = 40.3\%$; Q (3) = 5.03; p = .170). For reduction of symptoms ($k = 4$), the effect was also non-significant ($g = 0.01$ (95% CI: −0.11, 0.12; p = .93). Heterogeneity was low and non-significant ($I^2 = 0.0\%$; Q (3) = 0.52; p = .914). Not enough studies were available with mean number of completed sessions to make this comparison.

Moderator analyses were performed only for sufficiently large subgroups (viz. symptoms). For the moderators "anxiety disorder" and "mode of delivery", non-significant differences between the designated subgroups were found. The result for "reminders in both conditions" was marginally significant (Q (1) = 3.00; p = .08), showing an effect of $g = 0.00$ ($k = 6$) for those studies offering reminders in both groups, and $g = 0.24$ ($k = 4$) for those that did not. For the moderators "pre-treatment interview" and "CBT vs. other treatment" not enough studies were available in each subgroup to allow for meaningful comparisons. Most studies offered CBT and a pre-treatment interview.

3.5. Comparison 3: qualification of online counselors

Seven studies were found comparing different qualifications of online coaches/technicians and psychologist, or community-based vs. specialized psychologists): two studies from Hadjistavropoulos et al. (Hadjistavropoulos, Peynenburg, Nugent, et al., 2020; Hadjistavropoulos, Peynenburg, Thiessen, et al., 2020), Johnston et al. (2011), Robinson et al. (2010), Titov et al. (2009, 2010), and Westerhof et al. (2019). The pooled effect size was $g = 0.04$ (95% CI: −0.06, 0.14; p = .45), indicating that qualification was not associated with efficacy (Table 7). Heterogeneity was absent and non-significant ($I^2 = 0\%$; Q (6) = 5.79; p = .45). For adherence rates (Table 8), the weighed mean OR was 1.02 (95% CI: 0.27, 3.91; p = .97). Heterogeneity was absent and non-significant ($I^2 = 0\%$; Q (5) = 0.09; $p > .99$). Not enough studies were available that provided the average number of sessions completed for this analysis.

3.6. Publication bias

We inspected for the presence of publication bias in two sets of studies: those for human vs. technological guidance (k = 14) and those than compared degrees of human guidance (k = 10). Regarding the first set of studies reporting symptom outcomes, no signs of publication bias were present when inspecting the funnel plot for missing studies on the left. Using Duval and Tweedie's trim and fill procedure, no studies needed to be trimmed (random effects model). Likewise, Begg and Mazumdar's rank correlation test was non-significant ($\tau = -0.08$; p [one-tailed] = 0.35), as was the case for Egger's regression intercept (intercept = −0.40; p [one-tailed] = 0.18). For the subset of studies addressing degrees of human support (reporting symptoms), there was also no indication of publication bias. According to Duval and Tweedie's trim and fill procedure, no studies needed to be trimmed. Begg and Mazumdar's rank correlation test was non-significant ($\tau = 0.24$; p [one-tailed] = 0.16), as well as Egger's regression intercept (intercept = 0.86; p [one-tailed] = 0.17).

4. Discussion

This meta-analysis addressed the role of guidance in text-based internet treatments ("e-therapy") using a more fine-grained taxonomy of guidance than previous meta-analyses, and a broader range of mental complaints, while including only studies with direct comparisons between different types of guidance. Whilst previous meta-analyses usually focused on a categorical distinction of guidance versus self-help, we

Table 5
Results for degrees of human guidance – session average completed.

		95% Confidence interval				Heterogeneity			
Study name	Hedges'*g*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Berger2011*	−0.10	−0.65	0.44	−0.37	0.71				
Farrer2011*	0.26	−0.17	0.69	1.20	0.23				
Hadjistavropoulos2017*	0.51	0.21	0.81	3.30	*< 0.01*				
Hadjistavropoulos2020	0.05	−0.11	0.21	0.64	0.53				
Hadjistavropoulos2020b	0.07	−0.09	0.23	0.83	0.41				
*Oromendia2016**	*1.26*	*0.65*	*1.87*	*4.04*	*< 0.01*				
Pihlaja2020	0.48	0.08	0.87	2.36	0.02				
Weighed mean *g* (random effects)	**0.30**	**0.07**	**0.53**	**2.56**	**<0.05**	**24.13**	**6**	**<0.01**	**75%**
ITT only (k = 6)	0.27	0.03	0.52	2.16	<0.05				
Outliers removed (k = 6)	0.20	0.02	0.37	2.21	<0.05				

Note. Studies with * compare regular with optional support.
Study names in italics represent outliers.

Table 6
Results for degrees of human guidance – rate of patients completing all sessions.

		95% Confidence interval				Heterogeneity			
Study name	*OR*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Farrer2011*	1.15	0.36	3.68	0.24	0.81				
Gilbody2017*	2.07	1.03	4.19	2.03	0.04				
Hadjistavropoulos2017*	3.59	1.80	7.17	3.62	<0.01				
Hadjistavropoulos2020a	1.02	0.69	1.53	0.11	0.91				
Hadjistavropoulos2020b	1.12	0.78	1.62	0.61	0.54				
Kleiboer2015(1)*	1.10	0.62	1.95	0.31	0.75				
Klein2009	1.40	0.42	4.72	0.54	0.59				
Oromendia2016*	7.98	0.39	163.33	1.35	0.18				
Pihlaja2020	4.95	1.30	18.81	2.35	0.02				
Weighed mean *OR*	**1.57**	**1.09**	**2.25**	**2.44**	**<0.05**	**17.46**	**8**	**0.03**	**54%**
ITT only (k = 7)	1.54	1.00	2.38	1.96	<0.05				

Note. Studies with * compare regular with optional support.

Table 7
Results for qualification of therapists – symptoms.

		95% Confidence interval				Heterogeneity			
Study name	Hedges'*g*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Hadjistavropoulos2020a	0.11	−0.08	0.30	1.15	0.25				
Hadjistavropoulos2020b	0.01	−0.15	0.17	0.10	0.92				
Johnston2011	−0.34	−0.76	0.07	−1.62	0.11				
Robinson2010	0.10	−0.29	0.50	0.51	0.61				
Titov2010	0.09	−0.33	0.50	0.40	0.69				
Titov2009	0.03	−0.40	0.46	0.12	0.90				
Westerhof2019	0.54	−0.21	1.28	1.41	0.16				
Weighed mean *g* (random effects)	**0.04**	**−0.06**	**0.14**	**0.76**	**0.45**	**5.79**	**6**	**0.45**	**0%**

Table 8
Results for Qualification of therapists – rate of patients completing all sessions.

		95% Confidence interval				Heterogeneity			
Study name	*OR*	Lower limit	Upper limit	*Z*	p	*Q*	df	p	I^2
Hadjistavropoulos2020	0.99	0.04	26.71	−0.01	>0.99				
Johnston2011	0.96	0.04	24.21	−0.03	0.98				
Robinson2010	0.98	0.04	23.81	−0.01	>0.99				
Titov2009	1.10	0.05	25.76	0.06	0.95				
Titov2010	0.78	0.03	19.74	−0.15	0.88				
Westerhof2019	1.58	0.04	60.47	0.25	0.81				
Weighed mean *OR*	**1.02**	**0.27**	**3.91**	**0.04**	**0.97**	**0.09**	**5**	**>0.99**	**0%**

made two main comparisons: (1) human guidance vs. technological guidance and (2) degrees of intensity of human guidance. In addition, we compared studies with counselors of varying qualifications.

Our findings indicated that technological guidance was less efficacious compared to human guidance, which was found consistently across outcomes. For the purpose of our discussion, it is important to note that in most studies (11/14 = 79%), technological support was compared to regular (i.e., weekly) human support. Effects for symptoms and adherence were comparable. Yet, it is difficult to compare the two effect sizes directly, as they may differ in terms of their sensitivity to change. These effects also stem from slightly different subsets of studies, depending on availability of outcomes. In the studies included in this

review, technological support referred to basic support in the form of regular reminders in case of non-response, or motivating messages in case of response. According to our definition of technological support, this did not include continuous support with highly sophisticated techniques based on artificial intelligence (AI), such as chatbots (e.g., Bendig, Erb, Schulze-Thuesing, & Baumeister, 2019). Although technological advances are moving fast, the studies included in this review made use of basic, template-fixed messaging. These findings resonate with other meta-analytic findings (e.g., Riper et al., 2018), although this is the first meta-analysis based only on direct comparisons between technological and human support.

Results for varying degrees of human support were less consistent. Only for one of the outcome measures (adherence rates), results were significant and could be reliably estimated. This finding indicated that more intensive human support is more likely to reduce drop out than less intensive support. Furthermore, these effects turned out to be driven only by the subsets of studies comparing regular versus optional support. Although subgroups were small, these findings complement those of a recent meta-analysis focused on people with anxiety disorders, which included only two studies for this particular comparison (Domhardt et al., 2019). In practice optional guidance could be more intensive than regular guidance, this was not the case in the included studies. Some of the included studies in this comparison explicitly mentioned how often patients had initiated contact (Berger, Caspar, Richardson, Kneubühler, Sutter, et al., 2011, Hadjistavropoulos, Schneider, Edmonds, Karin, Nugentet, al., 2017, Kleiboer, Donker, Seekles, van Straten, Riper, et al., 2015; Oromendia, Orrego, Bonillo, & Molinuevo, 2016). In all of these studies, patients in the condition with support on demand ended up receiving less support than those with fixed (weekly) contact. To give a few examples: In one of the largest studies in this domain (Kleiboer et al., 2015), only 19% of those in the "support on request" condition asked for advice. Likewise, in the study by Berger et al. (2011), over half of the participants in the "Step-up condition" did not ask for additional support. In another study, patients in the "Optional support" condition logged in fewer times, spent fewer days enrolled in the program, and sent fewer and briefer emails to their therapists (Hadjistavropoulos et al., 2017). In conclusion, due to the small groups of studies, the finding that regular support yields better outcomes than optional support should be regarded as preliminary, and limited to adherence, and should be replicated with more primary studies.

In the current meta-analysis, we found an average attrition rate of 48% (range: 7–94%) for human guidance, and 51% (range: 26–86%) for technological guidance, which shows that the attrition is generally high in internet treatments, even with guidance. Compared to face-to-face CBT, for example, the dropout is approximately twice as high (24%; Linardon, Fitzsimmons-Craft, Brennan, Barillaro, & Wilfley, 2019). It is essential that therapists make an effort to increase adherence, and this meta-analysis indicates that offering human guidance aimed at therapeutic content slightly increases adherence, which in turn could increase efficacy. Future studies should aim to clarify whether staying in treatment indeed leads to better outcomes (mediation effect), which regular forms of therapy seem to indicate, or that patients drop out at the high-point of their optimal curve (Reich & Berman, 2020).

Because of the conflation between type of support and level of therapist qualification in previous studies (Dirkse et al., 2020; Johnston et al., 2011), obscuring clear conclusions, we analyzed studies comparing clinical to (mostly) non-clinical support separately under the heading of "qualification". We found that online counselors with higher levels of education/training were not more efficacious. This is in line with other meta-analyses, including partially overlapping studies (Baumeister et al., 2014; Domhardt et al., 2019), lending some support to the conclusion that online counselors with higher levels of training or education are not more efficacious. Yet, at the same time, in several studies included in these analyses, the more qualified counselor offered more support and/or moderated an online forum designed to assist a group of patients, instead of providing individual guidance (Robinson et al., 2010; Titov et al., 2009, 2010). These results should therefore be considered inconclusive, and future studies should strive to systematically disentangle qualification and intensity or format (group/individual) of treatment. This could also answer the question whether online group formats are more effective than individual formats.

4.1. Limitations

Some limitations need to be considered when interpreting the results. First, it should be kept in mind that for some of the included studies, the main aim was not to reduce mental distress, although this was our primary outcome measure. Some studies did not target depression or anxiety primarily, but addressed, for example, sleep difficulties or issues with eating instead. Effects for depression/anxiety for these studies may have been underestimated as a result but should have affected both tested treatment conditions in equal ways. Second, this paper focused on the immediate impact of subtleties of guidance and no conclusions can be drawn with respect to follow-up effects. Although there is burgeoning evidence that the effects of guidance may be long-lasting (Lancee et al., 2013; Oromendia et al., 2016; Ruwaard et al., 2009; Vernmark et al., 2010), more systematic support is needed. Third, our conclusions seem to apply mostly to the effects of online CBT, as these dominated the included studies. Furthermore, we tested for a differential impact of guidance on CBT versus other approaches, but other approaches were available to a limited extent, hampering sound conclusions. Fourth, not all authors were able to share data for adherence, which resulted in lower power for these meta-analyses. Moreover, although we did our best to obtain ITT samples, comparisons contained different degrees of completer and ITT samples, which could have resulted in inaccurate estimates. Fifth, we tested for the impact of therapist qualification on outcome, yet the variance in qualification was limited. Some of the studies compared two types of trained (specialized and community) psychologists, other studies compared one clinical psychologist to a person with no training, introducing a lot of "person-variance". We think that this needs further study, before any firm conclusions can be drawn. We did not consider cost-effectiveness in this study. Although we detected a small to moderate difference between regular vs. optional support, it remains to be considered whether the minor increases in efficacy outweigh the additional costs of more intensive human support. Please note that in case of regular support, support was scheduled at once a week, yet it was contingent upon participants completing their assignments. We could not systematically determine the actual frequency of support they received. Seventh, although we did our best to establish the degree of sophistication of the technological support in the studies included, we are not familiar with all programs used. Eighth, included publications were limited to the English language. On a related note, the included studies represent findings from Western, educated, industrialized, rich and democratic (WEIRD) countries. In the future, the questions raised in this meta-analysis should be extended to internet treatments in low and middle income countries, which were shown to be efficacious as well (Fu, Burger, Arjadi, & Bockting, 2020). Ninth, our ratings of methodological quality were based mostly on ratings by a junior psychologist, which could have limited their validity. Tenth, our moderator analyses were based on small subgroups, which likely resulted in low power to detect differences. These analyses should therefore be considered explorative. Eleventh, we excluded blended therapies from this meta-analysis, including those with videoconferencing. As a result of this, the generalizability of our findings to clinical practice may be reduced, as in clinical practice internet-based treatments are often provided in tandem with face-to-face services (e.g., Kooistra, Ruwaard, Wiersma, van Oppen, van der Vaart, et al., 2016; Wentzel, Van der, Bohlmeijer, & Van Gemert-Pijnen, 2016). Furthermore, most studies recruited in the community, which limits the generalizability of findings for clinical patients. Finally, we were not able to establish the impact of "technical support". We did not include this in our spectrum of guidance, because this type of

support is non-clinical. Most studies seem to offer this to participants, yet it was not systematically reported. Another type of support called "safety monitoring", was offered in some studies, and this is most pertinent to patients with depression and suicidal problems. It is possible that these types of support were offered "behind the scenes" in some studies, which could have obscured the impact of guidance that we tried to establish. In other words, participants in the minimal human support conditions may in some cases have received more support than we could reasonably determine.

Yet, despite these limitations, we believe our meta-analysis deepened and extended the knowledge concerning the impact of various levels and types of guidance on treatment outcomes for e-therapies, by introducing and employing a more nuanced and "clean" taxonomy of types of guidance. Most results can be considered robust and apply to a broad range of mental health difficulties. Moreover, most of the included studies had good to very good methodological quality.

4.2. Research implications

Some directions for future research should be considered. First, since we could not reach any definite conclusions concerning the qualification of online therapists, this should be more extensively studied. This could be done by including a sufficient number of highly qualified, experienced clinical psychologists or therapists, in comparison to, for instance, psychology students. Second, more systematic research is needed into the additional benefit of regular versus optional support. In this respect, it is important to be clear about any additional support that participants receive (e.g., safety monitoring, or technical support) to enable sound conclusions. Moreover, more detailed information is required concerning the degree of sophistication of the computer programs used, particularly with respect to (automated) reminders and motivating messages sent to the users. Taking these factors into account would allow to address additional, still more nuanced aspects of guidance, both human and non-human. Moving forward, it is likely that applications of e-therapy will become more sophisticated, and the impact of more interactive "conversational agents" should be incorporated into these examinations.

5. Conclusions

This meta-analysis indicates human support has superior effects over (simple) technological guidance alone. Findings regarding adherence further suggest that regular human guidance should be preferred over optional human guidance, but this finding did not generalize to clinical outcomes. These findings extend and refine previous reports addressing the coarse distinction between "guided" and "unguided" internet treatments. Future research should aim to clarify the added value of more qualified therapists, the degree to which regular over optional support is preferred, other questions related to mediation/moderation of efficacy, and the impact of more advanced use of AI-driven conversational agents.

Role of funding sources

This study was funded from internal resources from the University of Amsterdam.

Contributors

Jurrijn Koelen: Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Writing – Original Draft, Visualization, Supervision, Project administration **Anne Vonk**: Conceptualization, Methodology, Investigation **Anke Klein:** Writing - Review & Editing **Lisa de Koning**: Writing - Review & Editing **Peter Vonk**: Conceptualization, Writing - Review & Editing, Funding acquisition **Sabine de Vet**: Investigation, Formal Analysis **Reinout Wiers**: Conceptualization, Writing - Review & Editing, Funding acquisition

Declaration of Competing Interest

None.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2022.102179.

References[1]

Andersson, G. (2015). *The internet and CBT: a clinical guide*. Taylor & Francis.

Andersson, G., & Cuijpers, P. (2009). Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cognitive Behaviour Therapy, 38*(4), 196–205. https://doi.org/10.1080/16506070903318960

Andersson, G., & Titov, N. (2014). Advantages and limitations of internet-based interventions for common mental disorders. *World Psychiatry, 13*(1), 4–11. https://doi.org/10.1002/wps.20083

Andersson, G., Titov, N., Dear, B. F., Rozental, A., & Carlbring, P. (2019). Internet-delivered psychological treatments: from innovation to implementation. *World Psychiatry, 18*(1), 20–28. https://doi.org/10.1002/wps.20610

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

Begg, C. B., & Mazumdar, M.. Operating characteristics of a rank correlation test for publication bias Author (s): Colin B. Begg and Madhuchhanda Mazumdar Published by : International Biometric Society Stable. URL http://www.jstor.org/stable/2533446.

Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2019). Next Generation: Chatbots in Clinical Psychology and Psychotherapy to Foster Mental Health—A Scoping Review. *Verhaltenstherapie, 29*(4), 266–280. https://doi.org/10.1159/000499492

Berger, T., Caspar, F., Richardson, R., Kneubühler, B., Sutter, D., & Andersson, G. (2011). Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. *Behaviour Research and Therapy, 49*(3), 158–169. https://doi.org/10.1016/j.brat.2010.12.007

Berryhill, M. B., Culmer, N., Williams, N., Halli-Tierney, A., Betancourt, A., Roberts, H., & King, M. (2019). Videoconferencing Psychotherapy and Depression: A Systematic Review. *Telemedicine and E-Health, 25*(6), 435–446. https://doi.org/10.1089/tmj.2018.0058

Boettcher, J., Berger, T., & Renneberg, B. (2012). Does a pre-treatment diagnostic interview affect the outcome of internet-based self-help for social anxiety disorder? A randomized controlled trial. *Behavioural and Cognitive Psychotherapy, 40*(5), 513–528. https://doi.org/10.1017/S1352465812000501

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis (2.0)*. Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. United Kingdom: Wiley-Blackwell Publishing Ltd.

Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., & Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. *Cognitive Behaviour Therapy, 47*(1), 1–18. https://doi.org/10.1080/16506073.2017.1401115

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates. http://books.google.nl/books?id=Tl0N2lRAO9oC.

Cuijpers, P., Donker, T., Van Straten, A., Li, J., & Andersson, G. (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychological Medicine, 40*(12), 1943–1957. https://doi.org/10.1017/S0033291710000772

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). The role of common factors in psychotherapy outcomes. *Annual Review of Clinical Psychology, 15*(1). https://doi.org/10.1146/annurev-clinpsy-050718-095424

Deady, M., Choi, I., Calvo, R. A., Glozier, N., Christensen, H., & Harvey, S. B. (2017). eHealth interventions for the prevention of depression and anxiety in the general population: a systematic review and meta-analysis. *BMC Psychiatry, 17*(1), 1–14. https://doi.org/10.1186/s12888-017-1473-1

Dear, B. F., Staples, L. G., Terides, M. D., Fogliati, V. J., Sheehan, J., Johnston, L., … Titov, N. (2016). Transdiagnostic versus disorder-specific and clinician-guided versus self-guided internet-delivered treatment for Social Anxiety Disorder and comorbid disorders: A randomized controlled trial. *Journal of Anxiety Disorders, 42*, 30–44. https://doi.org/10.1016/j.janxdis.2016.05.004

Dirkse, D., Hadjistavropoulos, H. D., Alberts, N. A., Karin, E., Schneider, L. H., Titov, N., & Dear, B. F. (2020). Making Internet-delivered cognitive behaviour therapy scalable for cancer survivors: A randomized non-inferiority trial of self-guided and technician-guided therapy. *Journal of Cancer Survivorship, 14*(2), 211–225. https://doi.org/10.1007/s11764-019-00810-9

Domhardt, M., Geßlein, H., von Rezori, R. E., & Baumeister, H. (2019). Internet- and mobile-based interventions for anxiety disorders: a meta-analytic review of intervention components. *Depression and Anxiety, 36*(3), 213–224. https://doi.org/10.1002/da.22860

[1] See Supplement 2 for the references included in the meta-analysis.

Donkin, L., Christensen, H., Naismith, S. L., Neal, B., Hickie, I. B., & Glozier, N. (2011). A systematic review of the impact of adherence on the effectiveness of e-therapies. *Journal of Medical Internet Research, 13*(3), 1–12. https://doi.org/10.2196/jmir.1772

Duval, S., & Tweedie, R. (2000). A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634.

Farrand, P., & Woodford, J. (2013). Impact of support on the effectiveness of written cognitive behavioural self-help: a systematic review and meta-analysis of randomised controlled trials. *Clinical Psychology Review, 33*(1), 182–195. https://doi.org/10.1016/j.cpr.2012.11.001

Fu, Z., Burger, H., Arjadi, R., & Bockting, C. L. H. (2020). Effectiveness of digital psychological interventions for mental health problems in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet Psychiatry, 7*(10), 851–864. https://doi.org/10.1016/S2215-0366(20)30256-X

Hadjistavropoulos, H. D., Mehta, S., Wilhelms, A., Keough, M. T., & Sundström, C. (2020). A systematic review of internet-delivered cognitive behavior therapy for alcohol misuse: study characteristics, program content and outcomes. *Cognitive Behaviour Therapy, 49*(4), 327–346. https://doi.org/10.1080/16506073.2019.1663258

Hadjistavropoulos, H. D., Schneider, L. H., Edmonds, M., Karin, E., Nugent, M. N., Dirkse, D., … Titov, N. (2017). Randomized controlled trial of internet-delivered cognitive behaviour therapy comparing standard weekly versus optional weekly therapist support. *Journal of Anxiety Disorders, 52*, 15–24. https://doi.org/10.1016/j.janxdis.2017.09.006

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R. A Hands-on guide*. Chapmann & Hall/CRC Press.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Hedman, E., Furmark, T., Carlbring, P., Ljótsson, B., Rück, C., Lindefors, N., & Andersson, G. (2011). A 5-year follow-up of internet-based cognitive behavior therapy for social anxiety disorder. *Journal of Medical Internet Research, 13*(2). https://doi.org/10.2196/jmir.1776

Higgins, J. P., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions* (p. Version 5.1.0 (updated March 2011)). The Cochrane Collaboration.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Johansson, R., & Andersson, G. (2012). Internet-based psychological treatments for depression. *Expert Review of Neurotherapeutics, 12*(7), 861–870. https://doi.org/10.1586/ern.12.63

Johnston, L., Titov, N., Andrews, G., Spence, J., & Dear, B. F. (2011). A RCT of a Transdiagnostic Internet-Delivered Treatment for Three Anxiety Disorders: Examination of Support Roles and Disorder-Specific Outcomes. *PLOS ONE, 6*(11), Article e28079. https://doi.org/10.1371/journal.pone.0028079

Karyotaki, E., Efthimiou, O., Miguel, C., Bermpohl, F. M. G., Furukawa, T. A., Cuijpers, P., … Forsell, Y. (2021). Internet-based cognitive behavioral therapy for depression: A systematic review and individual patient data network Meta-analysis. *JAMA Psychiatry, 78*(4), 361–371. https://doi.org/10.1001/jamapsychiatry.2020.4364

Kleiboer, A., Donker, T., Seekles, W., van Straten, A., Riper, H., & Cuijpers, P. (2015). A randomized controlled trial on the role of support in Internet-based problem solving therapy for depression and anxiety. *Behaviour Research and Therapy, 72*, 63–71. https://doi.org/10.1016/j.brat.2015.06.013

Kocsis, J. H., Gerber, A. J., Milrod, B., Roose, S. P., Barber, J., Thase, M. E., … Leon, A. C. (2010). A new scale for assessing the quality of randomized clinical trials of psychotherapy. *Comprehensive Psychiatry, 51*(3), 319–324. https://doi.org/10.1016/j.comppsych.2009.07.001

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kooistra, L. C., Ruwaard, J., Wiersma, J. E., van Oppen, P., van der Vaart, R., van Gemert-Pijnen, … Riper, H. (2016). Development and initial evaluation of blended cognitive behavioural treatment for major depression in routine specialized mental health care. *Internet Interventions, 4*, 61–71. https://doi.org/10.1016/j.invent.2016.01.003

Linardon, J., Fitzsimmons-Craft, E. E., Brennan, L., Barillaro, M., & Wilfley, D. E. (2019). Dropout from interpersonal psychotherapy for mental health disorders: a systematic review and meta-analysis. *Psychotherapy Research : Journal of the Society for Psychotherapy Research, 29*(7), 870–881. https://doi.org/10.1080/10503307.2018.1497215

Lovell, K., & Richards, D. (2000). Multiple access points and levels of entry (maple): ensuring choice, accessibility and equity for cbt services. *Behavioural and Cognitive Psychotherapy, 28*(4), 379–391. https://doi.org/10.1017/s1352465800004070

Oromendia, P., Orrego, J., Bonillo, A., & Molinuevo, B. (2016). Internet-based self-help treatment for panic disorder: A randomized controlled trial comparing mandatory versus optional complementary psychological support. *Cognitive Behaviour Therapy, 45*(4), 270–286. https://doi.org/10.1080/16506073.2016.1163615

Reich, C. M., & Berman, J. S. (2020). Are psychotherapies with more dropouts less effective? *Psychotherapy Research, 30*(1), 23–40. https://doi.org/10.1080/10503307.2018.1534018

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: a systematic review and meta-analysis. *Clinical Psychology Review, 32*(4), 329–342. https://doi.org/10.1016/j.cpr.2012.02.004

Riper, H., & Cuijpers, P. (2016). Telepsychology and e-health. In J. C. Norcross, & D. Freedheim (Eds.), *vol 3. APA handbook of clinical psychology: Applications and methods* (pp. 451–463). American Psychological Association.

Riper, H., Hoogendoorn, A., Cuijpers, P., Karyotaki, E., Boumparis, N., Mira, A., … Smit, J. H. (2018). Effectiveness and treatment moderators of internet interventions for adult problem drinking: An individual patient data meta-analysis of 19 randomised controlled trials. *PLoS Medicine, 15*(12), 1–26. https://doi.org/10.1371/journal.pmed.1002714

Ruwaard, J., Schrieken, B., Schrijver, M., Broeksteeg, J., Dekker, J., Vermeulen, H., & Lange, A. (2009). Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. *Cognitive Behaviour Therapy, 38*(4), 206–221. https://doi.org/10.1080/16506070802408086

Salivar, E. J. G., Rothman, K., Roddy, M. K., & Doss, B. D. (2020). Relative Cost Effectiveness of In-Person and Internet Interventions for Relationship Distress. *Family Process, 59*(1), 66–80.doi. https://doi.org/10.1111/famp.12414

Spek, V., Cuijpers, P., Nykl\íček, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychological Medicine, 37*(3), 319–328. https://doi.org/10.1017/S0033291706008944

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *Journal of Clinical Epidemiology, 54*, 1046–1055.

Titov, N., Andrews, G., Schwencke, G., Solley, K., Johnston, L., & Robinson, E. (2009). An RCT Comparing Effect of Two Types of Support on Severity of Symptoms for People Completing Internet-Based Cognitive Behaviour Therapy for Social Phobia. *Australian & New Zealand Journal of Psychiatry, 43*(10), 920–926.

Titov, N., Andrews, G., Davies, M., Mcintyre, K., Robinson, E., & Solley, K. (2010). Internet treatment for depression: A randomized controlled trial comparing clinician vs. Technician assistance. *PLoS ONE, 5*(6), 1–9. https://doi.org/10.1371/journal.pone.0010939

Van't Hof, E., Cuijpers, P., & Stein, D. J. (2009). Self-help and internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. *CNS Spectrums, 14*(2 Suppl 3), 34–40. https://doi.org/10.1017/s1092852900027279

Vernmark, K., Lenndin, J., Bjärehed, J., Carlsson, M., Karlsson, J., Öberg, J., Carlbring, P., Eriksson, T., & Andersson, G. (2010). Internet administered guided self-help versus individualized e-mail therapy: a randomized trial of two versions of CBT for major depression. *Behaviour Research and Therapy, 48*(5), 368–376. https://doi.org/10.1016/j.brat.2010.01.005

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate* (2nd ed.). Routledge.

Weisel, K. K., Zarski, A. C., Berger, T., Krieger, T., Schaub, M. P., Moser, C. T., … Ebert, D. D. (2019). Efficacy and cost-effectiveness of guided and unguided internet- and mobile-based indicated transdiagnostic prevention of depression and anxiety (ICare Prevent): A three-armed randomized controlled trial in four European countries. *Internet Interventions, 16*, 52–64. https://doi.org/10.1016/j.invent.2018.04.002

Weiss, D. J., Nelson, A., Vargas-Ruiz, C. A., Gligorić, K., Bavadekar, S., Gabrilovich, E., … Gething, P. W. (2020). Global maps of travel time to healthcare facilities. *Nature Medicine, 26*(12), 1835–1838. https://doi.org/10.1038/s41591-020-1059-1

Wentzel, J., Van der Vaart, R., Bohlmeijer, E. T., & Van Gemert-Pijnen, J. E. W. C. (2016). Mixing online and face-to-face therapy: How to benefit from blended care in mental health care. *JMIR Mental Health, 3*(1), 1–7. https://doi.org/10.2196/mental.4534