

This is a repository copy of *Effectiveness and adherence of telephone-administered psychotherapy for depression : a systematic review and meta-analysis*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/150761/>

Version: Accepted Version

Article:

Castro, Adoración, Gili, Margalida, Ricci-Cabello, Ignacio et al. (5 more authors) (2020) Effectiveness and adherence of telephone-administered psychotherapy for depression : a systematic review and meta-analysis. *Journal of affective disorders*. pp. 514-526. ISSN 0165-0327

<https://doi.org/10.1016/j.jad.2019.09.023>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Effectiveness and adherence of telephone-administered psychotherapy for depression:

A systematic review and meta-analysis

Authors: Adoración Castro^{a,b}, Margalida Gili^{a,b,c}, Ignacio Ricci-Cabello^{b,d,e}, Miquel Roca^{a,b,c}, Simon Gilbody^f, Maria Ángeles Perez-Ara^{a,b}, Andrea Seguí^a, Dean McMillan^f

Affiliations:

^a Institut Universitari d'Investigació en Ciències de la Salut (IUNICS), University of Balearic Islands, Palma de Mallorca, Spain.

^b Institut d'Investigació Sanitaria Illes Balears (IdISBa), Mallorca, Spain

^c Primary Care Prevention and Health Promotion Research Network, RedIAPP, Madrid, Spain.

^d Gerencia de Atención Primaria Mallorca. IB-Salut. Mallorca, Spain

^e Ciber de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

^f Department of Health Sciences, University of York, York, United Kingdom.

Corresponding author:

Adoración Castro.
Institut Universitari d'Investigació en Ciències de la Salut (IUNICS).
University of Balearic Islands. Cra. de Valldemossa, km 7.5. Palma (Illes Balears).
Tel: +34 971 25 98 88 – E-mail: a.castro@uib.es

1. Introduction

Depression is one of the most prevalent mental health conditions, and is estimated to affect 320 million people worldwide (World Health Organization [WHO], 2017). Depressive disorders are associated with serious disability (van Schaik et al., 2004), loss in quality of life (Cuijpers et al., 2004), and substantial economic costs both at an individual and a societal level (Kessler, 2012; Smit et al., 2006).

Psychopharmacotherapy and psychotherapy are effective treatments for depression (Cuijpers et al., 2012; Cuijpers et al., 2008; Khan et al., 2012). Cognitive behaviour therapy (CBT), interpersonal psychotherapy (IPT) and behavioural activation (BA) have become main psychotherapy treatment options and are included in most guidelines as first-line treatment for depressive disorders (National Institute for Health and Clinical Excellence [NICE], 2009). Nevertheless, a number of barriers to traditional face-to-face psychotherapy, such as an insufficient number of trained professionals, its time consuming nature, the cost, and the perceived stigma of visiting a mental health professional, may prevent patients from accessing the available treatment (Bower & Gilbody, 2005; Brenes et al., 2011; Cuijpers et al., 2010; Kazdin & Blase, 2011; Kazdin & Rabbitt, 2013; Webb et al., 2017). Telephone-delivered psychotherapy can minimise such barriers (Brenes et al., 2011). Research has shown that telephone treatments can be cost-effective and has the potential to offer patients immediacy of help, anonymity and ease of access (Leach & Christensen, 2006; Mohr et al., 2008), are convenient for patients and therapists, eliminate treatment obstacles and can reduce treatment time by up to 40% (Lovell et al., 2006; Mohr et al., 2006).

Of the previous reviews on this area, one was published over a decade ago (Mohr et al., 2008) and a more recent review was limited to a narrative synthesis (Coughtrey & Pistrang, 2018). In addition, adherence to treatment data has not been systematically investigated before. Adherence to treatment is important to study because is considered a significant measure of acceptability, appropriateness and effect of a psychological treatment (van Ballegooijen et al., 2014). Indeed, for e-health therapies, adherence is associated with better depression outcomes (Donkin et al., 2011).

The present systematic review and meta-analysis aims to evaluate the effectiveness of telephone-administered psychotherapy for depression when compared to control conditions or other active treatments and to determine adherence to telephone-administered psychotherapy.

2. Methods

This systematic literature review was conducted according to the Centre for Reviews and Dissemination guidelines on conducting systematic reviews (CRD, 2009) and reported using the Preferred Reporting items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). The review protocol was registered with PROSPERO (registration number CRD42017076721).

2.1 Bibliographic searches

A range of databases (MEDLINE, Embase, PsycINFO and the Cochrane library) and grey literature sources (Open Grey Website, Conference Proceedings Citation Index in Web of Sciences (WoS) and Open Access Theses and Dissertations (OATD)) were examined. Trial registries (Clinicaltrials.gov and the International Clinical Trials Registry Platform) were also searched. All databases were examined from inception to 14 September 2017. The reference lists of all the included studies were also checked to identify further eligible studies. A bibliographical database was created using the website Covidence.org, which was used to store and manage the references.

The search terms, consisting of thesauri terms and free-text terms, were developed identifying search strategies of previous reviews and terms commonly used in potentially relevant studies identified in a scoping search. Search terms covered the constructs of “depression”, “psychotherapy” and “telephone”. The strategy was developed in MEDLINE (see appendix 1) and then adapted for the other databases. No restrictions were made in terms of language.

2.2 Inclusion and Exclusion Criteria

We included studies examining the impact of telephone-administered psychotherapy on depressive symptomatology in adults when compared to control or other active treatments. Specific eligibility criteria are detailed below: Participants: We included studies with adult participants (aged 18 and over) with major depression diagnosed using a structured clinical interview conducted according to internationally recognised standards (e.g., International Statistical Classification of Diseases and Related Health Problems (ICD), Diagnostic and Statistical Manual of Mental Disorders (DSM)) or significant (moderate to severe) depressive symptoms established using a validated screening measure (e.g., Patient Health Questionnaire (PHQ-9), Beck Depression Inventory (BDI)). Interventions: Studies were included if a treatment arm included telephone-administered of any kind of psychotherapy. Contact between therapist and patient had to be at least 90% over the telephone and the aim of the

intervention was to reduce depressive symptomatology. Comparator: Studies were included if the comparator was a control condition (e.g., waiting-list control, treatment as usual) or an active treatment (psychological or pharmacological). Outcomes: Studies were included if they measured the impact of the intervention on depression severity (the primary outcome of our review).

Study design: We only included randomised controlled trials (RCTs).

Context: Studies conducted in any setting were included. Restrictions were made in terms of publications (letter to editor, editorials, conference abstracts with no full text available). No restrictions were made in terms of language; however, one Korean-language study did not include because we could not obtain a translation of it.

2.3 Selection procedure

A standardised study eligibility form was used in the selection procedure based on the inclusion and exclusion criteria detailed above. Studies were evaluated for initial eligibility using title and abstracts against the study eligibility form. Full texts of the articles were searched for those studies that met the initial eligibility criteria and were examined once again using the eligibility form and a decision was made regarding their final inclusion in the review. All the references retrieved from the searches in the bibliographic databases were independently screened by two reviewers, with disagreements being solved by consensus with a third reviewer. Trial registries from Clinicaltrials.gov and reference lists of all included studies were screened by one reviewer.

2.4 Data extraction and Quality assessment

Data were extracted to a standardised data extraction form. The information extracted included study setting, sample characteristics (age, sex, ethnicity, presence of comorbidity and country), details of the intervention (type of the psychotherapy, number of sessions, session duration, frequency, other components of the intervention, and number of participants pre-intervention, post-intervention, and at follow-up), comparator group (type of comparator, and number of participants pre intervention, post intervention and at follow-up), adherence (mean number of sessions completed by the participants, percentage of patients who completed 100% of the intervention, percentage of patients who did not start the intervention), depression outcome measure, time points assessments and outcomes reported. Data on means, SDs, proportions and sample sizes were also extracted to permit calculation of effect sizes for each outcome.

We assessed the risk of bias of the included studies using the revised Cochrane risk-of-bias tool for randomized trials (Higgins et al., 2016). Items were scored as follows: low risk of bias, high risk of bias, or some concern, according to the Cochrane handbook descriptions. Two reviewers independently extracted all information and assessed the risk of bias. Disagreements were solved by consensus with a third reviewer.

2.5 Data analysis and synthesis

Two outcomes were considered in this review. The main outcome was the standardised measures of depression severity. Mean proportion of sessions completed by the participants (defined as mean number of sessions completed divided by the total number of sessions) was considered as a secondary measure. A narrative synthesis of the findings of included studies, involving a descriptive summary of each study, individual effect size and the quality appraisal was conducted.

If ≥ 2 studies were comparable in terms of the comparator (active treatment or control treatment), a meta-analysis was considered. We pooled data to summarise the difference in depression symptoms from baseline to the posttreatment scores between the intervention and comparator groups. We anticipated that the included trials would vary in their setting, intervention and design, so we used a random effects model to pool data (DerSimonian and Laird, 1986). The patient reported measures for depression varied between trials; so we used Hedge's method to calculate pooled effect sizes (Hedges, 1981) based on standardized mean difference. We standardized scores where required so that higher scores indicated higher levels of depression (Cohen, 1968). Where the standard deviation of the change between baseline and post-intervention was not provided, we derived them from baseline and final standard deviations, assuming a degree of correlation of 0.5 (Higgins, Deeks, & Altman, 2011). To assess the potential impact of this imputation method in the robustness of our findings, we conducted sensitivity analyses using a range of different correlations values (from 0.1 to 0.9). Heterogeneity was quantified by the I^2 statistic, where $I^2 > 50\%$ was considered evidence of substantial heterogeneity (Deeks & Higgins, 2011). Sources of heterogeneity were explored using Galbreith charts (Anzures-Cabrera & Higgins, 2010). When one or more studies were identified as being major contributors to a high level of heterogeneity (outliers), we removed them in a sensitivity analysis to assess the consistency of our findings. Publication bias was examined using funnel plots and the presence of asymmetry was assessed with Begg (Begg and Mazumdar, 1994) and Egger tests (Egger et al., 1997).

The preregistered review protocol specified subgroup and sensitivity analyses (Castro et al., 2017). Owing to the small number of studies with low risk of bias, it was not possible to conduct a sensitivity analysis based on quality assessment. Due to high heterogeneity, a sensitivity analysis was performed in which outlier studies were excluded. There were a sufficient number of studies to conduct the pre-specified subgroup analysis in the protocol in terms of clinical characteristics of the sample (depression vs. depression and long-term health conditions). Meta-analyses were conducted with STATA, version 12.0, using the command "metan". Subgroup differences according to the presence or absence of comorbidity were explored using the STATA option "by".

3. Results

3.1 Study selection

Search results are summarized in the PRISMA flow diagram (Figure 1). The initial search identified a total of 2,784 unique citations. Title and abstract screening of these citations resulted in the inclusion of 325 citations for further review. Following full text screening, 11 research reports (Alegría et al., 2014; Dwight-Johnson et al., 2011; Fann et al., 2015; Glueckauf et al., 2012; Himelhoch et al., 2013; Kalapatapu et al., 2014; Lam et al., 2013; Mohr et al., 2000; Mohr et al., 2012; Mohr et al., 2011; Piette et al., 2011) reporting on 10 separate trials, were finally included.

3.2 Characteristics of the included studies

A detailed description of characteristics of the included studies is provided in Table 1. The total combined sample size across the included trials was 1392 participants. All studies were conducted in the US except for one, which was conducted in Canada (Lam et al., 2013). All participants were adults and the majority were female (range: 9.5%-90.9%). Four studies reported on medical conditions as comorbidity with depression (Fann et al., 2015; Himelhoch et al., 2013; Mohr et al., 2000; Piette et al., 2011). One study was focused on employers (Lam et al., 2013), one in veterans (Mohr et al., 2011) and one study in primary care patients (Mohr et al., 2012).

All telephone interventions used CBT. Mean intervention duration was 15.7 weeks (range: 8-48) and mean number of treatment sessions was 12 (range: 6-21 sessions). Only one study included group sessions (Glueckauf et al., 2012); the remaining studies had individual sessions (Alegría et al., 2014; Dwight-Johnson et al., 2011; Fann et al., 2015; Himelhoch et al., 2013; Lam et al., 2013; Mohr et al., 2000, Mohr et al., 2012, Mohr et al., 2011; Piette et al., 2011).

Three studies compared telephone intervention to a control condition (Dwight-Johnson et al., 2011; Mohr et al., 2000, Mohr et al., 2011). Two studies compared telephone intervention to an active comparator (face-to-face CBT) and to a control condition (Fann et al., 2015; Alegría et al., 2014). Two studies compared telephone CBT to face-to-face CBT (Glueckauf et al., 2012; Mohr et al., 2012). Three studies compared the intervention condition to another active treatment: non-manualized face-to-face therapy (Himelhoch et al., 2013), self-help materials and a pedometer (Piette et al., 2011) and a prescription of antidepressant medication plus calls reminders (Lam et al., 2013).

3.3 Summary of findings

A detailed description of main results of included studies is provided in appendix 2. Three studies observed that telephone-administered psychotherapy was more effective than usual care in lowering depressive symptomatology (Fann et al., 2015; Alegría et al., 2014; Mohr et al., 2000). Nevertheless, two studies found no statistical differences between telephone-administered psychotherapy and usual care at post-treatment (Mohr et al., 2011; Dwight-Johnson et al., 2011). Four studies found no statistically significant differences in depression symptomatology when comparing telephone-administered psychotherapy to face-to-face psychotherapy (effect treatment condition) (Mohr et al., 2012; Himelhoch et al., 2013; Kalapatapu et al., 2014; Glueckauf et al., 2012) or other active comparator (Lam et al., 2013). One study found statistical significant differences between telephone-administered psychotherapy and an active comparator on depressive symptomatology, favouring telephone-administered psychotherapy (Piette et al., 2011).

Seven of 11 studies provide adherence outcome data and mean proportion ranges between 37% to 86% (Mohr et al., 2012; Dwight-Johnson et al., 2011; Fann et al., 2015; Glueckauf et al., 2012; Himelhoch et al., 2013; Piette et al., 2011; Lam et al., 2013).

3.4 Risk of bias

A summary of the revised Cochrane risk-of-bias tool for randomized trials can be found in Figure 2. Two trials presented low overall risk of bias (Himelhoch et al., 2013; Piette et al., 2011) and one some concerns (Lam et al., 2013). The remaining trials presented high risk of bias (Mohr et al., 2000; Mohr et al., 2011; Mohr et al., 2012; Dwight-Johnson et al., 2011; Fann et al., 2015; Glueckauf et al., 2012; Alegría et al., 2014). Most frequent biases were related to “measurement of the outcome domain” and “bias in selection of the reported result”. All trials presented low bias arising from the randomisation process.

3.5 Effectiveness of the interventions

The aim of the review was to evaluate the effectiveness of telephone-administered psychotherapy for depression when compared to control conditions or other active treatments at post-treatment and to determine adherence to telephone-administered psychotherapy.

3.5.1 Telephone-administered psychotherapy vs. control conditions

Four studies compared telephone-based intervention vs. a control group (Dwight-Johnson et al., 2011; Fann et al., 2015; Mohr et al., 2000; Mohr et al., 2011). The meta-analysis (available in Figure 3) showed that the telephone-administered psychotherapy produced larger reductions in depressive symptoms (standardized mean difference [SMD]= -0.85 (-1.56 to -0.15) when compared with control conditions. However, heterogeneity was high ($I^2=87.0\%$). This effect was robust (i.e., the differences remained statistically significant) to the use of a range of imputed correlation coefficients (which ranged from 0.1 to 0.9). A sensitivity analysis (see Figure 4) excluding a potential outlier (Dwight-Johnson et al., 2011) confirmed the statistically significant reduction in depressive symptoms, though with a smaller effect size (SMD= -0.48; 95% CI, -0.82, -0.14). Heterogeneity was substantially reduced ($I^2=23.5\%$).

3.5.2 Telephone-interventions versus Active Comparator

Six studies compared a telephone-based intervention with an active treatment: four studies compared telephone-based interventions to face-to-face interventions (Mohr et al., 2012; Fann et al., 2015; Glueckauf et al., 2012; Himelhoch et al., 2013), one study compared telephone-based intervention to antidepressant medication plus control reminders (Lam et al., 2013) and another compared telephone-based intervention with self-help and educational materials (Piette et al., 2011).

The meta-analysis (available in Figure 3) showed a non-significant effect size of -0.18 (SMD= -0.18, 95% CI -0.45 to 0.09); in favour of telephone-administered psychotherapy. Heterogeneity was high ($I^2=60.6\%$). This lack of statistically significant effect was consistent to the use of a range of imputed correlation coefficients. A sensitivity analysis (see Figure 5) removing two potential outliers (Himelhoch et al., 2013; Mohr et al., 2012) showed a significant larger impact on depressive symptoms favouring telephone-administered psychotherapy (SMD= -0.39, 95% CI, -0.58 to -0.21; $I^2= 0\%$).

There was no evidence of publication bias and asymmetry in any of the meta-analysis conducted.

3.5.3 Subgroup analysis according to the clinical characteristics of the sample

Subgroup analyses were performed for the two main comparisons according to the clinical characteristics of the sample (depression vs. depression and long-term health conditions (Table 2)).

Two studies compared telephone-administered psychotherapy vs. control conditions in patients with no comorbidities (Dwight-Johnson et al., 2010; Mohr et al., 2011), observing a large but not significant effect size of -1.00 in favour of telephone-administered psychotherapy (95% CI (-2.42 to 0.43)) and very high heterogeneity ($I^2=94.8\%$). An analysis removing potential outliers was not possible to perform due to small number of studies.

In patients with comorbidity, telephone-administered psychotherapy produced larger reductions in depressive symptoms when compared to control condition (SMD= -0.64; 95% CI= -1.12 to -0.15; $I^2=29.4\%$) (Fann et al., 2015; Mohr et al., 2000).

For studies comparing telephone-administered psychotherapy vs. active comparators in patients with no comorbidities (Glueckauf et al., 2012; Lam et al., 2013; Mohr et al., 2012), results showed a non-significant effect size of -0.0 (95% CI -0.27 to 0.13) and an absence of heterogeneity ($I^2= 0\%$).

In patients with comorbidity, three studies compared telephone-administered psychotherapy vs. active comparators (Fann et al., 2015; Himelhoch et al., 2013; Piette et al., 2010), observing a small but non-significant effect size of -0.18 in favour of telephone-administered psychotherapy (95% CI -0.72 to 0.34) and a high heterogeneity ($I^2= 72.2\%$).

3.6 Adherence to the telephone interventions

Seven studies (Mohr et al., 2012; Dwight-Johnson et al., 2011; Fann et al., 2015; Himelhoch et al., 2013; Piette et al., 2011; Lam et al., 2013, Glueckauf et al., 2012), reported sufficient data to estimate adherence to the telephone intervention (computed in terms of mean number of sessions completed by the participants divided by total number of intervention sessions).

The weighted average percentage of completed telephone sessions was 73% (range from 37% to 86%) (Table 3). A meta-regression analysis was planned a priori in order to explore the relationship between adherence and observed effect sizes. However, it was not possible to perform the analysis due to small number of studies that reported adherence outcome data.

Five studies provide data about the percentage of patients who completed the entire intervention (Mohr et al., 2000, Mohr et al., 2011, Mohr et al., 2012; Glueckauf et al., 2012; Fann et al., 2015). According to these studies, the number of participants completing treatment ranged from 68% and 85%. Three studies provided data on the percentage of participants who did not start the intervention: In Alegría et al., (2014), 10% of participants never started the intervention, in Dwight-Johnson et al., (2011), the figure was 12% and in Glueckauf et al., (2012), it was 14%.

4. Discussion

In this systematic review and meta-analyses we examined the effectiveness of telephone-administered psychotherapy for depression in adults when compared to control conditions or other active treatments and determined adherence treatment to telephone-administered psychotherapy. We identified 11 studies describing 10 separate trials observing that the majority of included studies had at least some methodological limitations.

We observed that telephone-administered psychotherapy produced a statistically significant effect on depressive symptomatology at post-treatment when compared to control conditions. Our result is consistent with findings from a previous meta-analysis conducted by Mohr et al. (2008). That review identified a total of 12 trials and the results showed a significant effect for telephone-delivered psychotherapy ($d= 0.26$), which is less than the effect size reported in our meta-analysis. Possible reasons for this discrepancy may be due to the control conditions used. Many of the control conditions used in studies included in that meta-analysis provided patients with active treatment conditions; whereas in our analysis, active and control conditions as comparators were analysed separately.

When a meta-analysis was performed for studies comparing telephone-administered psychotherapy to an active comparator, small and non-significant effects were observed (SMD=-0.18, 95% CI -0.45, 0.09). Our results are broadly in line with previous research. A meta-analysis conducted by Osenbach et al., 2013 found no differences between psychotherapy via synchronous telemental health (TMH) for depression (including telephone, internet and videoconferencing) vs. face-to-face psychotherapies. The same pattern was found in Bee et al. (2008). In their meta-analytic review, they included 13 studies and only two included in the review directly compared the efficacy of technology-mediated versus face-to-face psychotherapy for depression and anxiety, indicating a large but non-significant effect size of 0.55 in favour of technology-mediated psychotherapy. Despite the size effect reported here are small, our results are not far from the sorts of effects reported in previous meta-analyses

which examine the effectiveness of psychotherapies in Primary Care (Linde et al., 2015; Holvast et al., 2017; Wells et al., 2018), setting where depression is one of the more common problem. This is important to consider as it could be related to the external validity of our results.

When the analysis was rerun removing potential outliers, results showed that telephone-administered psychotherapy produced a statistically significant effect on depressive symptomatology. In this analysis, two studies directly compared face-to-face psychotherapy with telephone interventions (Fann et al., 2015; Glueckauf et al., 2012) and two used another less intense active comparator (Lam et al., 2013; Piette et al., 2011). Two of the studies having the most weight in the meta-analysis were the studies with less intense active treatment, which could lead to an overestimation of the effect of telephone treatment relative to other active treatments, suggesting that the telephone could be more effective than other less intense active treatments. This result suggests that the effects of the telephone-interventions for depression to active treatments are less conclusive than those studies which compare to control conditions. Moreover, in general one can also expect that effects of comparisons with active comparators are often smaller than effects with control comparators.

When a subgroup analysis according to clinical characteristics of the sample was performed, the greatest benefits were observed among the studies which compared telephone-administered psychotherapy to control conditions in patients with comorbidity (SMD= -0.64, 95% CI -1.12, -0.15). This finding are partially in line with a recent meta-analysis conducted by Park et al. (2018). That review included RCTs of telephone-delivered CBT for depression among patients with any chronic physical health condition. Results indicated that telephone-CBT had a significant effect on depression symptoms at post-treatment ($d=-0.20$, 95 CI: -0.29 – 0.10, $Z=4.06$, $p<.001$). The effect size reported in Park et al. (2018) is smaller than that reported here. This difference may be due to the fact that the included studies are different in both reviews in terms of comparator groups and participant characteristics. Despite these differences, both Park et al. (2018) and the current review indicate that telephone-CBT may be a useful intervention reducing barriers to treatment and improving depressive symptomatology for people with long-term physical health conditions (Park et al., 2018; Lovell 2010). The results from this review on this point should, however, be interpreted with caution, owing to the small number of studies included in the subgroup analyses and the presence of other possible differences between the chronic condition and non-chronic conditions studies apart from the presence of chronic health problems.

The second main objective of this systematic review and meta-analysis was to determine adherence to telephone-administered psychotherapy. Seven studies provided adherence outcome data (Mohr et al., 2012; Dwight-Johnson et al., 2011; Fann et al., 2015; Glueckauf et al., 2012; Himelhoch et al., 2013; Piette et al., 2011; Lam et al., 2013). Total weighted mean adherence was 73%. Our finding cannot be compared with other reviews of telephone-delivered treatment because this has not been reported in previous reviews, but it can be compared with other forms of treatment delivery. Van Ballegooijen et al. (2014) conducted a meta-analysis to analyse adherence to internet-based and face-to-face CBT for depression. They identified a total of 24 studies describing 26 treatment conditions (14 face-to-face CBT, 12 iCBT), with the following inclusion criteria: targeting depressed adults, no comorbid somatic disorder or substance abuse, community recruitment and published in the year 2000 or later up to 2013. Results indicated that patients completed on average 80% of the internet-based CBT. Similar results were found in face-to-face CBT. In this modality, the average percentage of completed sessions was 83%. Hence, results for telephone-administered psychotherapy seem to be broadly in line with other treatment modalities.

There were some limitations and strengths concerning the primary studies which should be mentioned. First, some of the included studies had a small sample size. Second, owing to variation in time points in the follow-up used in the studies, it was not possible to determine the effect of telephone-delivered psychotherapy in the medium and long term (e.g. 6 and 12 months post-treatment). Finally, a high risk of “bias in the measurement of the outcome” was found across nearly all studies because it is impossible to blind participants to whether they had received telephone-intervention. Therefore, a high bias in this domain is characteristic and almost inevitable for trials of psychological interventions. Strength of the primary studies is that all of them report adequate randomization process, and no baseline imbalances were found.

Some methodological limitations of the present systematic review should be taken into account when interpreting the results. Although no restrictions were made in terms of language, one Korean-language study was not included because we could not obtain a translation of it. The clinical heterogeneity in terms of the intervention (e.g., group vs. individual sessions), time received (in hours and sessions) from the treatment and the clinical population (e.g., depression, depression plus comorbidity) is also a limitation of the review. Finally, depression severity may be a moderator of treatment outcome, but we were unable to examine this in the current review. Similarly, the use of medication in addition to psychological

treatment may also influence treatment outcome, but we were not able to explore this in the review.

Implications and future research

Our findings suggest telephone-delivered psychotherapy is effective for the reduction of depressive symptomatology when compared to control conditions. However, there is no evidence that telephone-delivered psychotherapy treatment is less effective than active comparators. The finding that treatment delivered over the telephone is effective also has implications for online delivery of treatment, given that online delivery can be supplemented with telephone support. Furthermore, telephone-administered psychotherapy shows an adequate treatment adherence, similar to face-to-face and Internet interventions.

Nevertheless, available evidence is limited by some factors. The review identified a number of potential methodological biases in the primary studies that could be associated with an inflation of the effect sizes. To confirm the positive findings observed in our review, future research should include overcoming the methodological limitations of published work. Second, until now, only attrition rate were studied before. Nevertheless, the lack of a specific and common definition about treatment adherence and dropout at any point during studies causes disagreement and heterogeneity across the identified trials. A shared and specific definition about adherence and dropout is necessary in order to report a reliable rate and allowing us to make adequate comparisons across studies. Finally, additional research needs include an estimation of the cost-effectiveness of the interventions and an examination of their medium and long-term impact.

Funding

This work was supported by the PhD grant provided by the University of the Balearic Islands.

Declaration of interest: none

References

- Alegría, M., Ludman, E., Kafali, E. N., Lapatin, S., Vila, D., Shrout, P. E., . . . Canino, G., 2014. Effectiveness of the engagement and counseling for latinos (ECLA) intervention in low-income latinos. *Med. Care.* 52(11), 989-997.
- Anzures-Cabrera, J., Higgins, JP., 2010. Graphical displays for meta-analysis: An overview with suggestions for practice. *Res. Synth. Methods.* 1(1), 66-80.
- Bee, P. E., Bower, P., Lovell, K., Gilbody, S., Richards, D., Gask, L., & Roach, P., 2008. Psychotherapy mediated by remote communication technologies: A meta-analytic review. *BMC Psychiatry.* 8:60.
- Begg, C. B., Mazumdar, M., 1994. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 50(4), 1088-1101.
- Bower, P., Gilbody, S., 2005. Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. *Br. J. Psychiatry.* 186: 11-7.
- Brenes, G. A., Ingram, C. W., & Danhauer, S. C., 2011). Benefits and challenges of conducting psychotherapy by telephone. *Prof. Psychol. Resand Pr.* 42(6), 543-549.
- Castro, A., Gili, M., Roca, M., Ricci, I., Gilbody, S., McMillan, D., 2017. Effectiveness and adherence of telephone-administered psychotherapy for depression: a systematic review and meta-analysis.
http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017076721 (accessed 12 November 2018).
- Coughtrey, A.E., Pistrang, N., 2018. The effectiveness of telephone-delivered psychological therapies for depression and anxiety: A systematic review. *J. Telemed. Telecare.* 24(2), 65-74.
- Cuijpers, P., De Graaf, R., Van Dorsselaer, S., 2004. Minor depression: Risk profiles, functional disability, health care use and risk of developing major depression. *J. Affect. Disord.* 79(1-3), 71-79.
- Cuijpers, P., Reynolds, C. F., Donker, T., Li, J., Andersson, G., Beekman, A., 2012. Personalized treatment of adult depression: Medication, psychotherapy, or both? A systematic review. *Depres. Anxiety.* 29(10), 855-864.

- Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P., 2008. Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *J. Consult. Clin. Psychol.* 76(6), 909-922.
- Cuijpers, P., Van Straten, A., Warmerdam, L., & Van Rooy, M. J., 2010. Recruiting participants for interventions to prevent the onset of depressive disorders: Possible ways to increase participation rates. *BMC. Health. Serv. Res.* 10: 181.
- Deeks J, Higgins J., (2011). Chapter 9: Analysing Data and Undertaking Meta-analyses, in: Higgins, P. T., Green S., (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Accessed March 6, 2018 from: www.cochrane-handbook.org.
- DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. *Control. Clin. Trials.* 7(3), 177-88.
- 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. *J. Med. Internet. Res.* 13(3): e52.
- Dwight-Johnson, M., Aisenberg, E., Golinelli, D., Hong, S., O'Brien, M., & Ludman, E., 2011. Telephone-based cognitive-behavioral therapy for latino patients living in rural areas: A randomized pilot study. *Psychiatr. Serv.* 62(8), 936-942.
- Egger, M., Smith, G. D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 315(7109), 629-634.
- Fann, J. R., Bombardier, C. H., Vannoy, S., Dyer, J., Ludman, E., Dikmen, S., . . . Temkin, N., 2015. Telephone and in-person cognitive behavioral therapy for major depression after traumatic brain injury: A randomized controlled trial. *J. Neurotrauma*, 32(1), 45-57.
- Glueckauf, R. L., Davis, W. S., Willis, F., Sharma, D., Gustafson, D. J., Hayes, J., . . . Springer, J., 2012. Telephone-based, cognitive-behavioral therapy for african american dementia caregivers with depression: Initial findings. *Rehabil. Psychol.* 57(2), 124-139.
- Hedges, L. V., 1981. Distribution Theory for Glass's Estimator of Effect size and Related Estimators. *J. Educ. Behav. Stat.* 6(2), 107-128.
- Higgins, J.P.T., Deeks, J.J., Altman, D.G., (Eds.). Chapter 16: Special topics in statistics. In: Higgins, J.P.T., Green, S. (Eds.). *Cochrane Handbook for Systematic Reviews of Interventions*.

Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Accessed July 24 2019 from: www.cochrane-handbook.org.

Higgins, J.P.T., Sterne, J.A.C., Savović, J., Page, M.J., Hróbjartsson, A., Boutron, I., Reeves, B., Eldridge, S., 2016. A revised tool for assessing risk of bias in randomized trials In: Chandler J, McKenzie J, Boutron I, Welch V (eds). *Cochrane Methods*. Cochrane Database of Systematic Reviews, Issue 10 (Suppl 1).

Himelhoch, S., Medoff, D., Maxfield, J., Dihmes, S., Dixon, L., Robinson, C., . . . Mohr, D.C. 2013. Telephone based cognitive behavioral therapy targeting major depression among urban dwelling, low income people living with HIV/AIDS: Results of a randomized controlled trial. *AIDS. Behav.* 17(8), 2756-2764.

Holvast, F., Massoudi, B., Oude Voshaar, R.C., & Verhaak, P.F.M. 2017. Non-pharmacological treatment for depressed older patients in primary care: A systematic review and meta-analysis. *PLoS One.* 12(9), e0184666

Kalapatapu, R. K., Ho, J., Cai, X., Vinogradov, S., Batki, S. L., Mohr, D. C., 2014. Cognitive-behavioral therapy in depressed primary care patients with co-occurring problematic alcohol use: Effect of telephone-administered vs. face-to-face treatment-A secondary analysis. *J. Psychoactive Drugs*, 46(2), 85-92.

Kazdin, A. E., Blase, S. L., 2011. Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspect. Psychol. Sci.* 6(1), 21-37.

Kazdin, A. E., Rabbitt, S. M., 2013. Novel models for delivering mental health services and reducing the burdens of mental illness. *Clinical Psychological Science*, 1(2), 170-191.

Kessler, R. C., 2012. The costs of depression. *Psychiatr. Clin. North Am.* 35(1), 1-14.

Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., & Brown, W. A., 2012. A systematic review of comparative efficacy of treatments and controls for depression. *PLoS ONE*, 7(7)

Lam, R. W., Parikh, S. V., Ramasubbu, R., Michalak, E. E., Tam, E. M., Axler, A., . . . Manjunath, C. V., 2013. Effects of combined pharmacotherapy and psychotherapy for improving work functioning in major depressive disorder. *Br. J. Psychiatry.* 203(5), 358-365.

Leach, L. S., Christensen, H., 2006. A systematic review of telephone-based interventions for mental disorders. *J. Telemed. Telecare.* 12(3), 122-129.

- Linde, K., Sigterman, K., Kriston, L., Rücker, G., Jamil, S., Meissner, K., & Schneider A., 2015. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. *Ann. Fam. Med.* 13(1), 56-68.
- Lovell, K., Cox, D., Haddock, G., Jones, C., Raines, D., Garvey, R., Roberts, C., Hadley, S., 2006. Telephone administered cognitive behaviour therapy for treatment of obsessive compulsive disorder: Randomised controlled non-inferiority trial. *BMJ.* 333(7574), 883-886.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* 6(7), e1000097.
- Mohr, D. C., Carmody, T., Erickson, L., Jin, L., Leader, J., 2011. Telephone-administered cognitive behavioral therapy for veterans served by community-based outpatient clinics. *J. Consult. Clin. Psychol.* 79(2), 261-5.
- Mohr, D. C., Hart, S. L., & Marmar, C., 2006. Telephone administered cognitive-behavioral therapy for the treatment of depression in a rural primary care clinic. *Cogn. Ther. Res.* 30(1), 29-37.
- Mohr, D. C., Ho, J., Duffecy, J., Reifler, D., Sokol, L., Burns, M. N., Jin, L., Siddique, J., 2012. Effect of telephone-administered vs face-to-face cognitive behavioral therapy on adherence to therapy and depression outcomes among primary care patients: A randomized trial. *JAMA.* 307(21), 2278-2285.
- Mohr, D. C., Likosky, W., Bertagnolli, A., Goodkin, D. E., Van Der Wende, J., Dwyer, P., & Dick, L. P., 2000. Telephone-administered cognitive-behavioral therapy for the treatment of depressive symptoms in multiple sclerosis. *J. Consult. Clin. Psychol.* 68(2), 356-61.
- Mohr, D. C., Vella, L., Hart, S., Heckman, T., & Simon, G., 2008. The effect of telephone-administered psychotherapy on symptoms of depression and attrition: A meta-analysis. *Clin. Psychol.* 15(3), 243-253.
- National Institute for Health and Clinical Excellence (NICE), 2009. Depression in Adults: The treatment and management of depression in adults. Accessed Jan 29, 2018, from <https://www.nice.org.uk/guidance/cg90/chapter/1-Guidance#stepped-care>.
- Osenbach, J., O'Brien, K., Mishkind, M., Smolenski, D., 2013. Synchronous telehealth technologies in psychotherapy for depression: a meta-analysis. *Depress. Anxiety.* 30(11), 1058-67.

Park, S., Kim, Y., & Oh, E., 2018. Effectiveness of Telephone-Delivered Cognitive Behavioral Therapy for Depression in Patients with Chronic Physical Health Conditions: A Meta-Analysis. *J. Korean. Acad. Psychiatr. Ment. Health. Nurs.* 27(3), 227-239.

Piette, J. D., Richardson, C., Himle, J., Duffy, S., Torres, T., Vogel, M., Barber, K., & Valenstein, M., 2011. A randomized trial of telephonic counseling plus walking for depressed diabetes patients. *Med. Care*, 49(7), 641-648.

Smit, F., Cuijpers, P., Oostenbrink, J., Batelaan, N., de Graaf, R., & Beekman, A. (2006). Costs of nine common mental disorders: Implications for curative and preventive psychiatry. *J. Ment. Health. Policy. Econ.* 9(4), 193-200.

University of York, 2009. CRD. Systematic reviews: Centre for reviews and dissemination's guidance for undertaking reviews in health care. Accessed April, 15 2017, from https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.

Van Ballegooijen, W., Cuijpers, P., Van Straten, A., Karyotaki, E., Andersson, G., Smit, J. H., & Riper, H., 2014. Adherence to internet-based and face-to-face cognitive behavioural therapy for depression: A meta-analysis. *PLoS ONE*, 9(7), e100674.

Van Schaik, D. J. F., Klijn, A. F. J., Van Hout, H. P. J., Van Marwijk, H. W. J., Beekman, A. T. F., De Haan, M., & Van Dyck, R., 2004. Patients' preferences in the treatment of depressive disorder in primary care. *Gen. Hosp. Psychiatry*, 26(3), 184-189.

Webb, C. A., Rosso, I. M., & Rauch, S. L., 2017. Internet-based cognitive-behavioral therapy for depression: Current progress and future directions. *Harv Rev Psychiatry*. 25(3), 114-122.

Wells, M.J., Owen, J.J., McCray, L.W., Bishop, L.B., Eells, T.D., Brown, G.K., ... Wright, J.H. 2018. Computer-Assisted Cognitive-Behavior Therapy for Depression in Primary Care: Systematic Review and Meta-Analysis. *Prim. Care. Companion. CNS. Disord*, 20(2). pii: 17r02196. World Health Organization (WHO), 2017. Depression and Other Common Mental Disorders, accessed Sept, 28 2018 from <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf?sequence=1>

Figure 1: Flowchart of articles included at each stage of the screening process

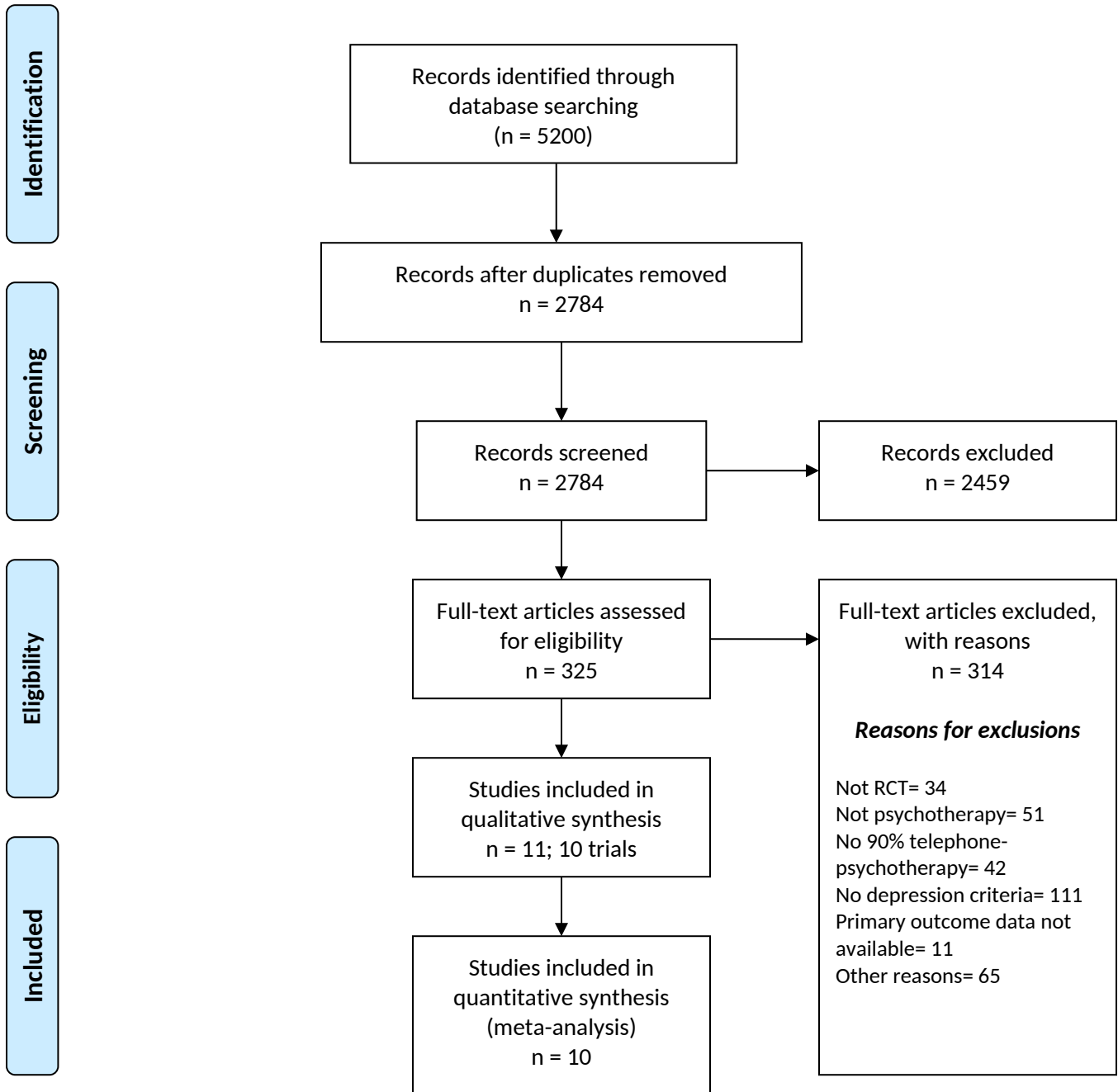


Figure 2:
Quality Assessment: The revised Cochrane
Risk of Bias Tool (RoB 2.0)

Key:	
	Low risk of bias
	High risk of bias
	Some concerns risk of bias





























































	Bias arising from the randomization process	Bias due to deviations from the intended interventions	Bias due to missing outcome data	Bias in measurement of the outcome	Bias in selection of the reported result	Overall bias
Alegria et al., 2014						
Dwight-Johnson et al., 2011						
Fann et al., 2015						
Glueckauf et al., 2012						
Himmelhoch et al., 2013						
Lam et al., 2013						
Mohr et al., 2000						
Mohr et al., 2011						
Mohr et al., 2012						
Piette et al., 2010						

Figure 3: Effect of telephone-administered psychotherapy on depressive symptoms in comparison to active treatments and control conditions.

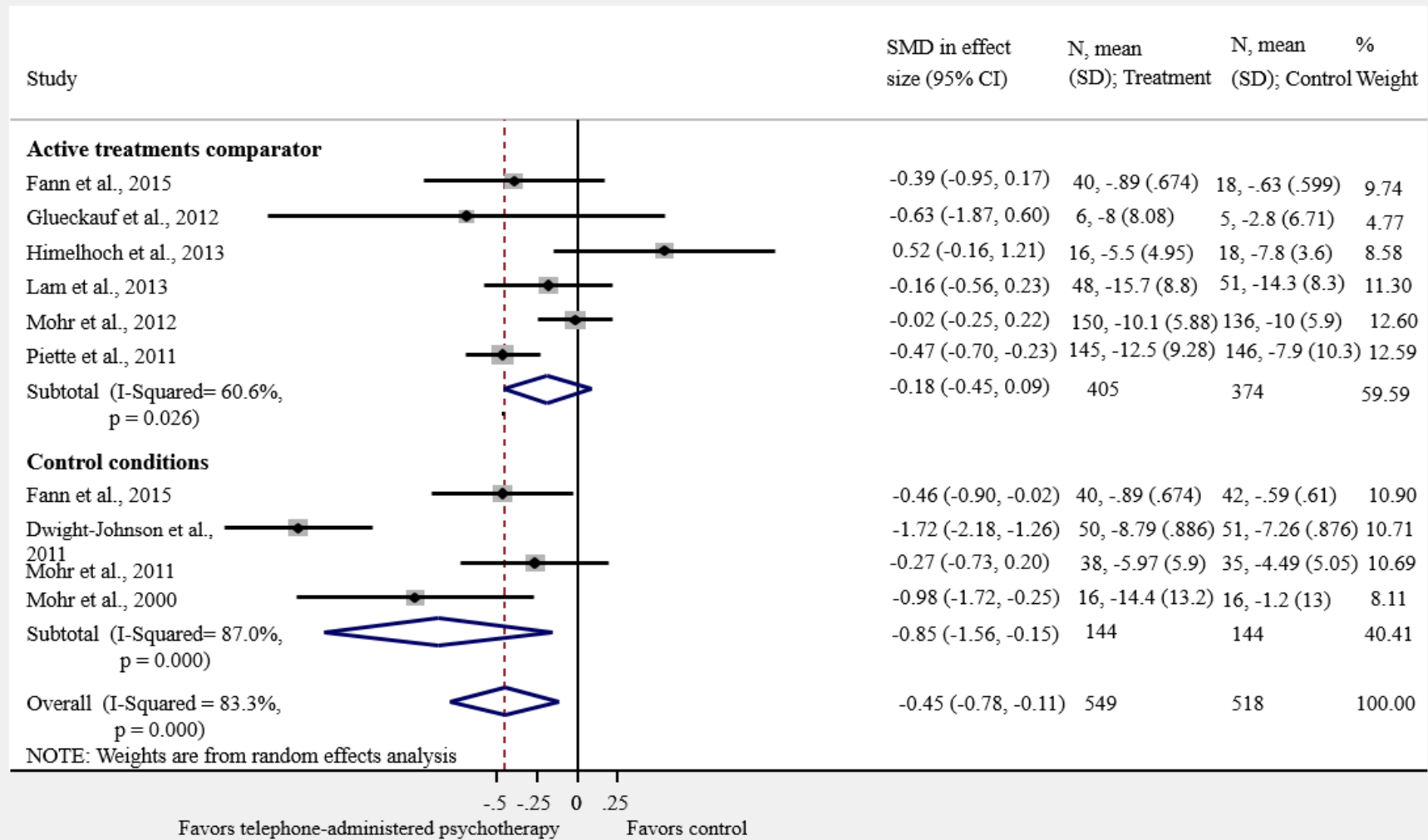


Figure 4: Sensitivity analysis showing effects of telephone-administered psychotherapy on depressive symptoms in comparison to control conditions.

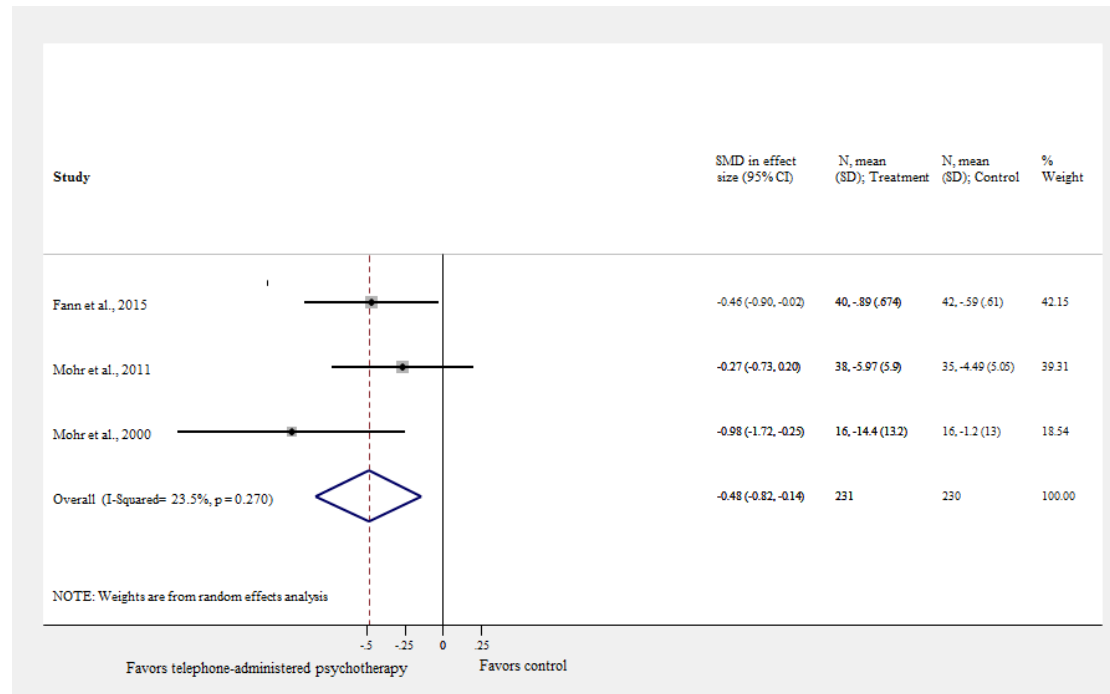


Figure 5: Sensitivity analysis showing effects of telephone-administered psychotherapy on depressive symptoms in comparison to active treatments.

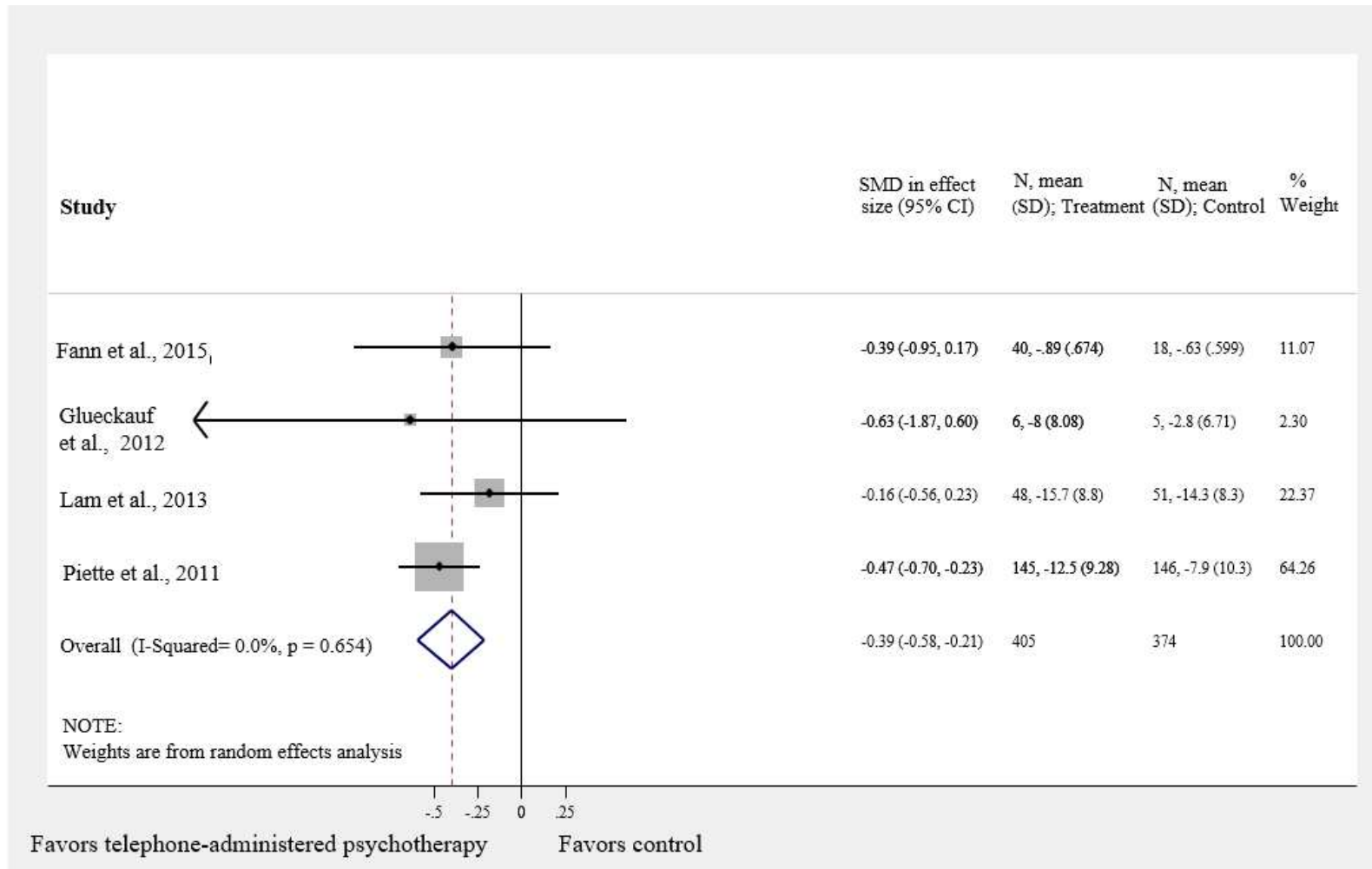


Table 1a: Bibliographic searches – search strategy (Medline)

Medline (Ovid SP)		
Searches		Results
1	Depression/	101163
2	exp Depressive Disorder/	100577
3	depressed.ti,ab.	85443
4	depression.ti,ab.	253891
5	depressive.ti,ab.	84766
6	1 or 2 or 3 or 4 or 5	382737
7	Psychotherapy/	52646
8	exp behavior therapy/	67759
9	exp emotion-focused therapy/	2
10	exp gestalt therapy/	164
11	exp narrative therapy/	127
12	exp person-centered therapy/	416
13	exp psychoanalytic therapy/	16990
14	exp psychotherapy, brief/	3340
15	exp psychotherapy, multiple/	714
16	exp psychotherapy, psychodynamic/	373
17	exp psychotherapy, rational-emotive/	196
18	exp Cognitive Therapy/	23578
19	exp counseling/ not exp pastoral care/ not exp sex counseling/	36734
20	psychotherap*.ti,ab.	35169
21	cognitive therap*.ti,ab.	2340
22	cognitive behavio*.ti,ab.	19248
23	CBT.ti,ab.	6956
24	behavio* therap*.ti,ab.	15541
25	behavio* activation.ti,ab.	1373
26	BA.ti,ab. not 25	19354
27	interpersonal psycho*.ti,ab.	860
28	counseling.ti,ab.	50695
29	counselling.ti,ab.	20452
30	collaborative care.ti,ab.	1484
31	7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30	258310
32	exp Telephone/	20997
33	exp Telemedicine/	22742
34	exp Telenursing/	178
35	telephon*.ti,ab.	48850
36	phoning.ti,ab.	52
37	calls.ti,ab.	27749
38	callback*.ti,ab.	117
39	call* back*.ti,ab.	381
40	cellphone.ti,ab.	107
41	mobilephone.ti,ab.	1
42	((mobile or cell) and phone).ti,ab.	4694
43	32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42	105566
44	6 and 31 and 43	1202

exp= explode (all trees); .ti,ab= search in title and abstract

Table 2a: Main results of the studies identified

Reference	Primary outcome: Depressive symptomatology	Secondary outcome: Adherence data
Alegría et al., 2014	For the PHQ-9, ECLA-T ($p=0.01$; $ES=.64$) was significantly better than usual care in lowering depressive symptoms.	Average number of sessions completed by the participants was not reported.
Dwight-Johnson et al., 2011	No statistical differences was found between T-CBT and UC in depressive symptomatology at posttreatment assessment (T-CBT: mean=8.23 vs. UC: mean= 10.08; $p=.165$).	Patients in the T-CBT group completed on average 4.62 ($SD=3.19$) sessions of a total of 8 sessions.
Fann et al., 2015	The CBT-T group had significantly more improvement on the SCL-20 than the UC group (treatment effect = 0.36, 95% CI: 0.01- 0.70; $p = 0.043$).	Participants in the T-CBT completed on average 9.6 ($SD= 3.3$) telephone sessions of a total of 12 sessions
Glueckauf et al., 2012	No statistically significant effects were obtained for group (telephone vs. f-to-f CBT) and the group X time interaction ($p>.05$) in depression.	On average, participants in the T-CBT completed a 10.28 sessions of a total of 12 sessions.
Himelhoch et al., 2013	There were no statistically significant differences on depression treatment outcomes comparing face-to-face psychotherapy to T-CBT whether evaluating outcomes on the QID-SR (9.2 ± 3.7 vs. 10.8 ± 5.5 ; $p = 0.28$).	On average, participants attended 4.1 ($SD=2.7$) sessions of a total of 11 sessions.
Lam et al., 2013	There were no significant between-group differences in MADRS score ($d=0.16$).	The mean number of T-CBT sessions completed by the participants was 6.4 ($SD= 2.8$) of a total of 8 sessions.
Mohr et al., 2000	The telephone-psychotherapy group showed lower levels of depressive symptoms than the UCC group at posttreatment ($p = .03$).	Average number of sessions completed by the participants was not reported.
Mohr et al., 2011	There were no significant Time X Treatment effects for the PHQ-9, $F(1, 157) = 1.64$, $d = 0.37$, $p = .20$.	Average number of sessions completed by the participants was not reported.
Mohr et al., 2012	There were no significant treatment differences at posttreatment between T-CBT and face-to-face CBT on the PHQ-9 ($P=.89$). The intention-to-treat posttreatment effect size on the PHQ-9 it was $d=-0.02$ (90% CI,-0.20 to 0.17). Both results were within the inferiority margin of $d=0.41$, indicating that T-CBT was not inferior to face-to-face CBT.	The mean number of T-CBT sessions completed by the participants was 15.5 ($SD= 4.4$) of a total of 18 sessions.

Kalapatapu et al., 2014	No statistical differences between face-to-face CBT and T-CBT groups on depressive outcomes at posttreatment assessment (5.9 (5.4) vs. 6.9 (7.2), $W_s=2,074.5$; $z= 0.04$, $P=0.97$ on PHQ).	The mean number of T-CBT sessions completed by the participants was 14.7 (SD=5.2) of a total of 18 sessions.
Piette et al., 2011	Both intervention and control groups experienced a significant ($p<.05$) improvement in their average BDI depression scores, with a 4.54 point greater average improvement in the intervention than control group ($p<.0001$)	On average, intervention patients completed 13.5 out of a possible 21 telephone CBT sessions.

ECLA-T: Effectiveness of the Engagement and Counseling for Latinos-Telephone; ES: Effect size; QIDS: Quick Inventory of Depressive Symptomatology; MADRS: Montgomery-Asberg Depression Rating Scale; PHQ-9: Patient Health Questionnaire 9 items; SCL-20: Hopkins Symptom Checklist Depression Scale 20 items; SD: Standard Deviation; T-CBT: Telephone Cognitive-Behavioral Therapy; CBT: Cognitive-Behavioral Therapy; UC: Usual Care; 95% CI: 95% Confidence Interval.

Table 1: Characteristics of the studies included

Reference	Setting, Sample and country	Intervention	Comparator	Depression instrument measure and Time Points
Alegría et al., 2014	Setting: Community-based clinics. Sample: low-income Latinos. Age (%): 35-49 yr: 37% Percentage of women: 82%. US and Puerto Rico	Telephone-Based CBT 4 individual weekly sessions + 2 individual biweekly sessions + Workbook and CBT exercises. N=87.	UC: N=86; Face-to-Face Psychotherapy (same as intervention but delivered face-to-face). N=84.	PHQ-9; Measurement time points: Baseline, posttreatment, 4 mo FU.
Dwight-Johnson et al., 2011	Setting: Primary Care Sample: Latinos living in rural areas. Age (M [SD]): Overall 39.81 (10.56). Percentage of women: 78%. US	Telephone-Based CBT 8 x 50 min individual weekly sessions + Patient Workbook. N=50	Enhanced UC: N=51.	PHQ-9; Measurement time points: Baseline, midpoint, posttreatment and 6 mo.
Fann et al., 2015	Setting: Community and clinical settings. Sample: Traumatic Brain Injury patients. Age (M [SD]): Overall 45.8 (13.3). Percentage of women: 37%.	Telephone-CBT 12 x 30-60 min individual weekly sessions + Workbook. N=40	In-Person CBT (same as intervention but delivered face-to-face). N= 18; UC: N=42.	SCL-20; Measurement time points: Baseline, interim assessment, posttreatment and 6 mo.
Glueckauf et al., 2012	Setting: Community and clinical setting.	Telephone-Based CBT: 12 (7 group and 5	Face-to-Face CBT (same as intervention but delivered in	CES-D; Measurement time points: Baseline and posttreatment.

	Sample: African American Dementia Caregivers. Age (M [SD]) ^o : Overall 58.09 (10.11). Percentage of women: 90.9%	individual sessions) x 60 min weekly + CBT guidebook, a copy of <i>The 36 Hour day</i> and information about local dementia care resources prior to the first training session. N=7	person) N=7;	
Himelhoch et al., 2013	Setting: HIV clinics. Sample: HIV patients. Age (M [SD]): Overall 45.12 (8.33). Percentage of women: 73.5%	Telephone-based CBT: 11 x 45 min individual sessions Weekly + Workbook and a linked therapist manual. N=16.	Non-Manualized face-to-face CBT: 11 x 60-min blocks. N=18.	QIDS; Measurement time points: Baseline, midpoint and posttreatment.
Lam et al., 2013	Setting: Community and clinical settings. Sample: employers. Age (M [SD]) ^o : Intervention: 42.3 (10.4). Comparator: 44.2 (9.9). Percentage of women: 54.54%	Telephone CBT: 8 x 30-40 min individual sessions Weekly + escitalopram 10/20 mg/day. N=52;	Active condition: 10-minute structured telephone call weekly, enquiry about progress and reminders to take medication + escitalopram 10/20 mg/day N=53.	MADRS; Measurement time points: baseline, midpoints and posttreatment
Mohr et al., 2000	Setting: Hospital care. Sample: multiple sclerosis. Age (M [SD]): Intervention: 42.6 (12.8) Comparator: 42.1 (9.4). Percentage of women: 71.87%	Telephone-administered CBT: 8 x 50 min individual weekly sessions + Patient workbook. N=16.	UC: N= 16.	POMS; Measurement time points: Baseline and posttreatment.

Mohr et al., 2011	Setting: CBOCs. Sample: veterans. Age (M [SD]): Overall 55.9 (10.59). Percentage of women: 9.4%	Telephone-Administered CBT: 16 x 45-50 min individual weekly sessions + Patient workbook. N=41.	TAU: N= 44.	PHQ-9; Measurement time points: Baseline, midpoint (12weeks), posttreatment and 6 mo FU.
Mohr et al., 2012	Setting: Primary care. Sample: Age (M [SD]): Intervention: 47.8 (12.6). Comparator: 47.5 (13.5). Percentage of women: 77.55%.	Telephone-Administered CBT: 18 x 45 min individual sessions 2 weekly sessions, 12 weekly and 2 booster sessions during 4 weeks + Patient workbook. N=163.	Face-to-face CBT (same as intervention but delivered in person); N=162.	PHQ-9; Measurement time points: Baseline, midpoints, posttreatment (week 18), 3 mo follow up and 6mo FU.
Kalapatapu et al., 2014*	Setting: Primary Care. Sample: problematic alcohol use diagnosis. Age (M [SD]): Intervention: 45.6 (13.7) Comparator: 41.9 (13.9) Percentage of women: 87.35%	Telephone-Administered CBT: 18 x 45 min individual sessions 2 sessions weekly, 12 weekly and 2 booster sessions during 4 weeks + Patient workbook. N=50	Face-to-face CBT (same as intervention but delivered in person); N=53.	PHQ-9; Measurement time points: Baseline, midpoints, posttreatment, 3 mo FU.
Piette et al., 2011	Setting: Community and clinical settings. Sample: diabetes patients. Age (M [SD]) ^o :	Telephone-delivered CBT program: 21 x (min NR) individual sessions	Active condition: patients received a copy of the Feeling Good Handbook -a self-help CBT book for depression, NIMH	BDI; Measurement time points: Baseline and posttreatment.

	Overall: 56 (10.1). Percentage of women: 51.5%	12-weekly and 9 monthly booster sessions + Patient Manual and pedometer N=172.	educational depression materials, walking and diabetes educational materials and a list of local resources for depression + pedometer. N= 167.	
--	--	--	---	--

CBT: Cognitive-Behavioral Therapy; UC: Usual Care; TAU: Treatment As Usual; ECLA-T: Effectiveness of the Engagement and Counseling for Latinos-Telephone; NIMH: National Institute of Mental Health; PHQ-9: Patient Health Questionnaire 9 items; SCL-20: Hopkins Symptom Checklist Depression Scale 20 items; CES-D: Center for Epidemiological Studies Depression; QIDS: Quick Inventory of Depressive Symptomatology; MADRS: Montgomery-Asberg Depression Rating Scale; POMS: Profile of Mood States; BDI: Beck Depression Inventory; Mo: months; FU: posttreatment follow-up; US: United States; *Secondary analysis of Mohr et al., 2012. ^o based on data included in the analysis.

Table 2. Subgroup meta-analysis according to the clinical characteristics of the sample.

Meta-analysis And Subgroup	References	SMD in effect size (95% CI)	Statistical Heterogeneity I² (%)
Telephone- administered psychotherapy vs. control condition			
Depression alone	Dwight-Johnson et al., 2011; Mohr et al., 2011	-1.00 (-2.42, 0.43)	94.8%
Depression and physical comorbid condition	Fann et al., 2011; Mohr et al., 2000	-0.64 (-1.12, -0.15)*	29.4%
Telephone- administered psychotherapy vs. active comparator			
Depression alone	Glueckauf et al., 2012; Lam et al., 2013; Mohr et al., 2012	-0.00 (-0.27, 0.13)	0%
Depression and physical comorbid condition	Fann et al., 2015; Himelhoch et al., 2013; Piette et al., 2010	-0.18 (-0.72, 0.34)	72.2%

SMD: Standardized mean difference. *Favors telephone-administered psychotherapy over comparator.

Table 3. Secondary outcome: Adherence data

Reference	N randomized to telephone-administered psychotherapy	Mean number of sessions completed (SD)/Total number of intervention sessions	Mean Proportion
Alegría et al., 2014	87	NR/6	NA
Dwight-Johnson et al., 2011	50	4.62 (3.19)/8	0.57
Fann et al., 2015	40	9.6 (3.3)/12	0.8
Glueckauf et al., 2012	7	10.28 (NR)/12	0.85
Himmelhoch et al., 2013	16	4.1 (2.7)/11	0.37
Lam et al., 2013	52	6.4 (2.8)/8	0.8
Mohr et al., 2000	16	NR/8	NA
Mohr et al., 2011	41	NR/16	NA
Mohr et al., 2012	163	15.5 (4.4)/18	0.86
Piette et al., 2011	172	13.5 (NR)/21	0.64
Total weighted average proportion			0.7284

SD= Standardized Deviation; NR= Not reported; NA: Not applicable.

Declaration of interest: none

Author Statement

Contributors

A Castro, D McMillan and S Gilbody designed the review and wrote the protocol. A Castro and D McMillan developed the search strategy. A Castro undertook the bibliographic searches. I Ricci-Cabello, MA Pérez-Ara, A Seguí and A Castro screened citations, extracted the relevant data and assessed risk of bias. I Ricci-Cabello undertook the statistical analysis. A Castro, M Gili and M Roca wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Role of the Funding Source

This work was supported by the PhD grant provided by the University of the Balearic Islands.

Acknowledgements

The authors would like to thank all the authors of the cited papers for getting back to them about any queries.