Effects of psychological treatment for depression among people not actively seeking help: a meta-analysis

Ruiying Zhao[1] , Arpana Amarnath[1], Eirini Karyotaki[1,2] , Sascha Y. Struijs[1,3] and Pim Cuijpers[1,2]

[1]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [2]WHO Collaborating Centre for Research and Dissemination of Psychological Interventions, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands and [3]Institute of Clinical Psychology, Leiden University, Leiden, The Netherlands

Abstract

Although psychological treatments have been found to be effective for depression in adults, many individuals with depression do not actively seek help. It is currently unclear whether psychological treatments are effective among those not actively seeking help. Besides, little is known about the proportion of patients who completed a screening questionnaire who end up in a clinical trial. Therefore, we conducted a meta-analysis of 52 randomized trials comparing psychotherapies for adults with a diagnosis or elevated symptoms of depression against control conditions (care-as-usual, waiting list, and other inactive treatment). Only studies recruiting participants who do not actively seek help (participants who have been recruited through screening instead of advertisements and clinical referrals) were included. To obtain an overall effect estimate of psychotherapy, we pooled all post-test differences with a random-effects model. We found that psychological treatments had a moderate to high effect on reducing depressive symptoms compared to control groups [$g = 0.55$; 95% confidence interval (CI) 0.41–0.69]. Heterogeneity was high ($I^2 = 75\%$; 95% CI 68–80). At 12 months' follow-up, the effects were small but significant (6–8 months: $g = 0.33$; 95% CI 0.14–0.52; 9–12 months: $g = 0.24$; 95% CI 0.11–0.37). As a secondary outcome, we found that 13% of patients who completed a screening questionnaire met the inclusion criteria for depression and agreed to be randomized in the trial. Based on the current evidence, psychological treatments for depression might be effective for depressed patients who are not actively seeking help.

CAMBRIDGE UNIVERSITY PRESS

Introduction

Major depression is the most prevalent mental health disorder worldwide. The World Health Organization estimates that 322 million people (4.4% of the global population) suffer from depression (WHO, 2017). Depression is the leading cause of disability worldwide and is associated with considerable mortality and morbidity (Cuijpers et al., 2014b; Vos et al., 2016). Not surprisingly, it has a debilitating impact on individuals' daily functioning, quality of life, and wellbeing. In addition, depression also poses a substantial burden on society, including high levels of service use, enormous economic costs, and production losses (Berto, D'Ilario, Ruffo, Virgilio, & Rizzo, 2000; Friedrich, 2017; Greenberg & Birnbaum, 2005).

Given depression's detrimental effect on individuals and society, hundreds of randomized controlled trials (RCTs) have been conducted in the past four decades. The results demonstrated that adult depression can be treated effectively with psychological treatments (Cuijpers et al., 2021b; Cuijpers, Karyotaki, de Wit, & Ebert, 2020). However, many people with depression do not actively seek help, and the treatment uptake is low in the general population. For instance, in the USA, only 35.3% of individuals with severe depression reported having seen a mental health professional in the previous year (Pratt & Brody, 2014). The low help-seeking rate is associated with several barriers, such as the limited availability of professional resources, financial considerations, low perceived need, mental illness stigma, lack of time, preference of self-management, and skepticism about treatment effectiveness (Andrade et al., 2014; Cuijpers, 2021; Mojtabai, 2009; Schnyder, Panczak, Groth, & Schultze-Lutter, 2017).

Considering the critical nature of this situation, several strategies have been suggested to increase depression help-seeking, including education and awareness campaigns targeting individual attitudes and mental health literacy (Salerno, 2016), but also training for primary care providers (Lipson, Speer, Brunwasser, Hahn, & Eisenberg, 2014). Further, tailored interventions to increase help-seeking intentions and behaviors (Ebert et al., 2019; Xu et al., 2018),

and indirect prevention and treatment (Cuijpers, 2021) may also help enhance depression help-seeking. Besides, systematic screening may be another possible solution to help in reaching people not actively seeking help and providing timely treatment. For instance, it has been suggested that programs combining early recognition of depression with an integrated system for management, result in reduction or remission of depression (Siu et al., 2016). However, several other systematic reviews found that evidence for routine screening is not very strong (Keshavarz et al., 2013; Meijer et al., 2013; Thombs et al., 2013; Thombs, Ziegelstein, Roseman, Kloda, & Ioannidis, 2014). And therefore, some national preventive services policy groups recommend against screening programs (NICE, 2009; Thombs, Markham, Rice, & Ziegelstein, 2021). Therefore, it is currently controversial whether screening programs should be recommended to enhance the low uptake of treatments.

Despite this lack of consensus, a fundamental question that should be considered before screening can be recommended is if we offer treatments to those who do not actively seek help, what will happen, and whether the treatments would still be effective. A systematic review, which included six studies, investigated the effects of psychotherapy for people who do not actively seek help in primary care (Cuijpers, van Straten, van Schaik, & Andersson, 2009). That meta-analysis indicated that when patients in primary care who do not seek help were offered treatments, there was no significant difference between psychotherapy and controls [waiting list, care-as-usual (CAU), and other control groups] [$k = 6$, $d = 0.13$; 95% confidence interval (CI) −0.08 to 0.34]. However, another study found that compared with usual care, screening with adequate treatment for adult depression in primary care could improve outcomes (Pignone et al., 2002). Evidently, whether psychological treatments are effective when people are not actively seeking help so far remains unclear. In addition, the number of included trials was too small to draw a definite conclusion. Furthermore, those studies were only focused on primary care, and many new trials in primary care and other settings have been conducted since then.

Therefore, we decided to conduct a meta-analysis to examine the effects of psychological treatments of depression among adults in any setting not actively seeking help but who were approached by researchers and screened for trial eligibility. In addition to the main analyses of intervention effects, we explored the proportions of participants that were included in the trial compared to the total number of potential participants who completed the screening scales for depression. By doing this, we hope to know when actively identifying people who do not actively seek help in the trial, how many of them should be screened to get enough patients.

Methods

The present meta-analysis has been preregistered at the Open Science Framework (https://osf.io/8sf6m) and was reported in accordance with The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Moher et al., 2009). The Amendments to the protocol can be seen in online Supplementary Appendix A.

Eligibility criteria

For the current study, we employed the following inclusion criteria: (1) patient: studies recruiting participants who do not actively seek help (we defined 'people who do not actively seek

help' as participants who have been recruited through screening instead of advertisements and clinical referrals). (2) Outcome: the primary focus was treating adults with elevated symptoms of depression above a cut-off score on a validated self-report scale or unipolar depression based on a clinical diagnostic interview. (3) Intervention and comparison: studies comparing psychological treatment with a control condition, i.e. CAU, waiting list, and other inactive treatment. (4) Studies employed an RCT design. We excluded studies that recruited participants through advertisements and clinical referrals, and studies conducted in specialized mental health care.

Identification and selection of studies

We used an established database of studies examining psychological treatments for adult depression. This database has been described in detail elsewhere (Cuijpers et al., 2020). For this database, we conducted an ongoing comprehensive search in PubMed, PsycINFO, Embase, and the Cochrane Library (last update: 1st January 2021). The full search string can be found in online Supplementary Appendix B. Moreover, we searched reference lists of earlier meta-analyses on psychological treatments for depression. Titles and abstracts from the search were independently screened by two reviewers. Studies considered as potentially relevant according to one of the researchers were retrieved as full text. For this specific paper, based on the existing database, two reviewers (R.Z. and A.A.) independently screened the full texts to find relevant studies. Disagreements were solved by seeking the opinion of a third, senior reviewer (P.C.).

Data extraction and classification

Data were independently extracted by two reviewers. Discrepancies were discussed and checked by revisiting the original paper. We coded the following information:

(1) Participant characteristics: age group; percentage of female participants; mean age; diagnosis/self-report as an inclusion criterion.
(2) Psychotherapy information: type [was defined according to the generic definitions of therapies given in Cuijpers et al. (2020), treatment format (individual, group, guided self-help, telephone, or other formats], number of sessions.
(3) General characteristics of the studies: type of control group; setting for screening; year of publication; the country where a study was conducted; western country (yes/no).
(4) Information regarding assessment tools.
(5) Outcomes: data on calculating effect sizes (number of participants in intervention and control groups, means and standard deviations for both groups at post-test); data on long-term outcomes (more than 6 months); data on study drop-out (for any reason) for each of the psychotherapy and control conditions.
(6) The number of patients who received the screening scales for depression, the number of patients that completed screening tools, and the number of patients who were randomized in the trial.

Risk of bias assessment

We used the first version of the Cochrane Risk of Bias Assessment Tool to assess the validity of the included studies (Higgins et al.,

2011). The assessments were conducted in the following domains: random sequence generation; concealment of allocation to conditions; blinding of assessors or use of self-report scales; appropriate methods for handling incomplete outcome data (which was evaluated as high risk of bias when intention-to-treat analyses were not used and/or it had overall more than 50% study dropout and/or more than 30% imbalance in missing outcomes between groups); and selective outcome reporting (rated as positive when prospectively registered primary outcomes were consistently reported in the article) (Miguel, Karyotaki, Cuijpers, & Cristea, 2021). Each domain with a lack of information was rated as high risk. Following this, studies met at least four quality criteria were assessed as overall low risk of bias. Two researchers (R.Z. and A.A.) assessed these domains independently, and disagreements were resolved through discussion until consensus was reached.

Outcome measures

Our primary outcome was the effect size of each comparison between psychotherapy and a control group, indicating the difference between two groups at post-test assessment. We calculated the effect size (Hedges' *g*) based on mean scores, standard deviations, and the number of participants. As many included studies had small sample sizes, we used Hedges' *g*, which is an effect size that adjusts for bias in small sample studies (Olkin & Hedges, 1985). The effect sizes of 0.2, 0.5, and 0.8 represent small, moderate, and large effect sizes, respectively. If means or standard deviations were not available, we used the procedures of Comprehensive Meta-Analysis software (CMA, version 3.3.070) to calculate the effect size using dichotomous outcomes or other statistics (such as *t* value, *p* value, or change scores). If a study (or a comparison between psychotherapy and a control group) included more than one depression assessment, they were initially pooled within the study or comparison before pooling across studies, so each study or comparison contributed one effect size to the overall analysis.

Meta-analyses

Data analysis (except for the calculation of effect sizes, which was done in CMA) was performed in R studio for Windows (version 1.3.959), by using the 'meta' (Balduzzi, Rücker, & Schwarzer, 2019), 'metafor' (Viechtbauer, 2010), and 'dmetar' (Harrer, Cuijpers, Furukawa, & Ebert, 2022) packages in R (version 4.0.1). As we expected considerable heterogeneity among the studies, we chose a random-effects pooling model in all analyses.

We pooled all included comparisons to obtain an overall effect estimate of psychotherapy. We also calculated numbers-needed-to-treat (NNT) to explain the clinical meaning of the effect size, by using the formulae provided by Furukawa (1999). We conservatively set the event rate in the control group as 19% (according to the pooled response rate of 50% reduction of symptoms across trials in psychotherapy for depression) (Cuijpers et al., 2014a). Heterogeneity was examined with I^2 and its 95% CI. In general, heterogeneity can be considered as low (25%), moderate (50%), and substantial (75%) (Higgins, Thompson, Deeks, & Altman, 2003).

A series of sensitivity analyses were conducted to examine the effect: (1) after the exclusion of outliers (the 95% CI of the effect size did not overlap with the 95% CI of the pooled effect size); (2) for studies with low risk of bias only; (3) for studies included multiple comparisons (two sensitivity analyses in which only one

effect size per study was included: one with the smallest effect size and one with the largest effect size).

Furthermore, subgroup analyses were conducted to test whether the effects are different based on the characteristics of the study (setting, age group, diagnosis at baseline, and risk of bias). All subgroup analyses were performed according to the mixed-effects model, in which studies within subgroups were pooled with the random-effects model, whereas tests for significant differences between subgroups were conducted with the fixed-effects model. We also conducted multivariate meta-regression analyses to explore potential predictors. The outcome variable was the effect size, and the predictors were setting, age group, diagnosis, and risk of bias.

Additionally, funnel plots and Egger's test were used to examine potential publication bias (Egger, Smith, Schneider, & Minder, 1997). When we found indications of publication bias, Duval and Tweedie's trim-and-fill procedure was implemented to adjust for possible bias, through detecting missing studies and imputing them (Duval & Tweedie, 2000).

Finally, study dropout (for any reason) was calculated with the relative risk (RR). We also examined long-term effects for studies with more than 6-month follow-up outcomes since randomization.

Additional analyses

As a secondary outcome, we calculated the pooled proportion of participants randomized in the trial from the total number of potential participants who completed the screening scales for depression. We also examined the pooled proportion of participants randomized in the trial from the total number of people who received the screening scales for depression.

We conducted two separate meta-analyses using 'metaprop' command of the 'meta' package in R (version 4.0.1). In these analyses, the logit transformation was first applied to each study's proportion. After pooling binomial outcome data, we back-transformed the synthesized results to the raw proportion scale. Since considerable heterogeneity was expected, we adopted a random-effects model. The between-study component of variance τ^2 was estimated by use of the DerSimonian–Laird method.

Results

Selection and inclusion of studies

Our search strategy identified a total of 27 133 records. After removing duplicates, we screened 19 612 records based on titles and abstracts. A total of 3239 studies were retrieved for full-text review, and 3187 studies were excluded from further investigation. The PRISMA flow chart of the study selection process and reasons for exclusion are presented in Fig. 1. As a result, 52 studies (with 61 comparisons between a psychotherapy and a control group) fulfilled all inclusion criteria and were included in the meta-analyses. References are presented in online Supplementary Appendix C.

Characteristics of included studies

The key characteristics of the included studies are summarized in Table 1. In sum, 7755 participants were involved in the meta-analysis (intervention groups: *N* = 3788; control groups: *N* = 3967). The sample size ranged from 19 to 570, with a mean age range from 21.9 to
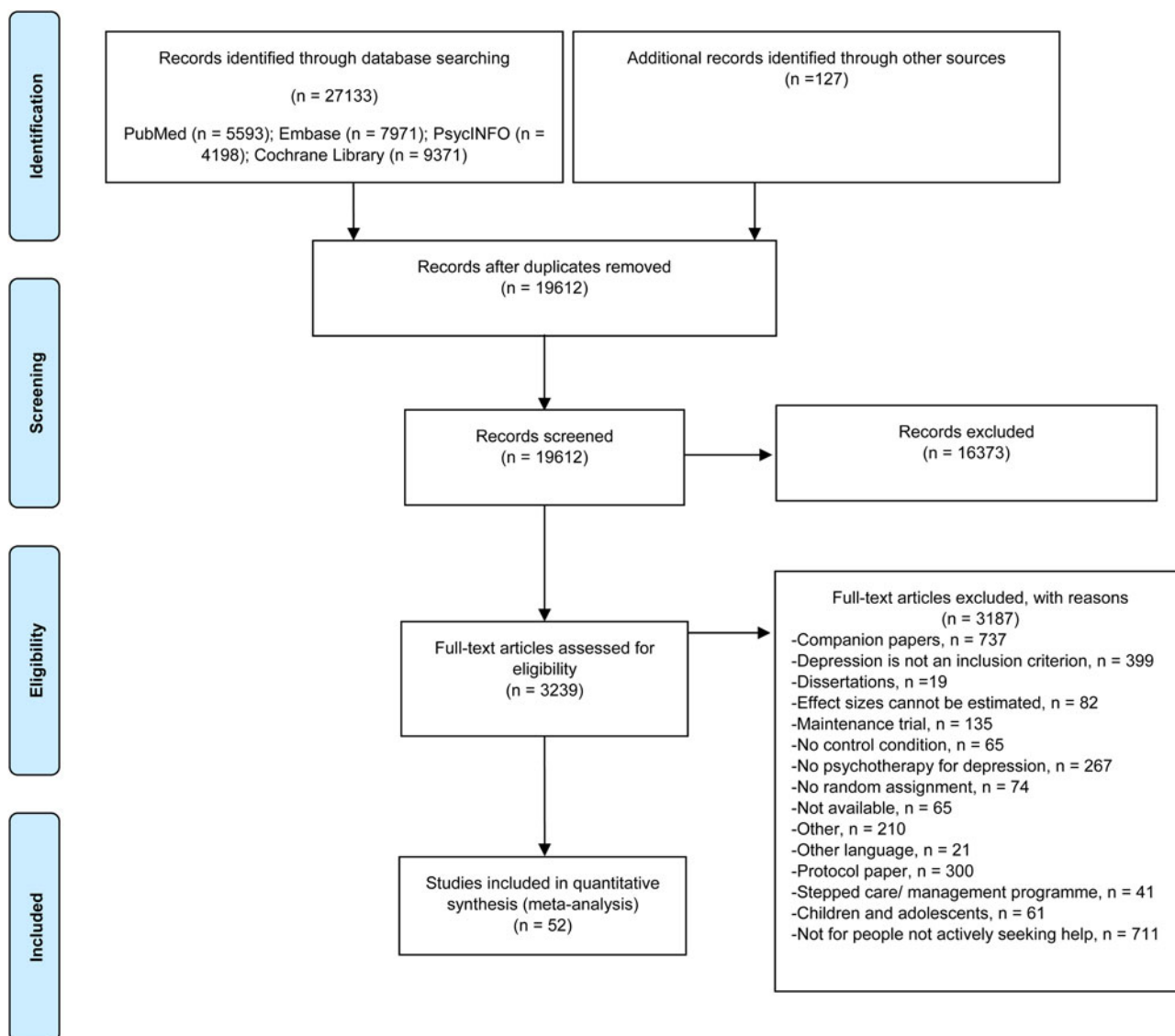
Fig. 1. Flowchart on the selection of studies.

81.5. Studies were conducted in general medical care ($n = 22$; 42.3%), perinatal care ($n = 14$; 26.9%), primary care ($n = 5$; 9.6%), and another context ($n = 11$; 21.2%). Additionally, 52 studies resulted in 52 proportions of participants in the trial from the number of participants who completed the screening tools. The proportions of participants in the trial from the number of participants who received the screening tools were available for 47 studies (online Supplementary Appendix D). A comprehensive description of characteristics can be seen in online Supplementary Appendix E.

Quality assessment

Overall, the risk of bias was considerable. Of the 52 studies, 39 reported an adequate sequence generation (75%); 31 reported allocation to conditions by an independent party (59.6%); 51 reported blinding of outcome assessors or used only self-report outcomes (98.1%); and 36 used appropriate methods to handle missing data (69.2%). Besides, 10 studies (19.2%) were rated as free from selective reporting. In total, 26 studies (50.0%) that

met at least four quality criteria received a low risk of bias. The remaining half of the included studies met less than four criteria were rated as a high risk of bias.

Overall effects of psychotherapies for depression

As shown in Fig. 2 and Table 2, the overall pooled effect size of 61 comparisons between psychotherapies and control conditions was 0.55 (95% CI 0.41–0.69), which corresponds with an NNT of 5.46. Heterogeneity was high ($I^2 = 75\%$; 95% CI 68–80).

Exclusion of nine outliers led to a somewhat smaller effect size ($g = 0.47$; 95% CI 0.39–0.55; NNT = 6.64), and to a considerable reduction of heterogeneity ($I^2 = 30\%$; 95% CI 2–51). Limiting the analyses to studies with low risk of bias did not affect much to the effect size ($g = 0.57$; 95% CI 0.33–0.82; NNT = 5.25), and the level of heterogeneity ($I^2 = 82\%$; 95% CI 75–87).

Because more than one psychotherapy was compared with the same control group in seven studies (five had two psychotherapy arms, and two had three arms), we conducted two sensitivity

Table 1. Summary of key characteristics of the included studies

Study	g	s.e.	Therapy	Ctr	Settings	Age group	M age	Pr wom	Dx	Frm	Nsess	Country	SG	AC	BA	IOD	SOR
Ammerman, 2013a	0.87	0.22	cbtc	cau	ppd	Adults	21.90	1.00	+	ind	11	North A	−	+	+	+	−
Beutel, 2014	0.52	0.19	dyn	cau	genmed	Adults	51.80	1.00	+	ind	18	EU	+	+	+	+	−
Burns, 2007	0.40	0.22	other	cau	genmed	Elderly	81.10	0.77	−	ind	6	UK	+	+	+	+	−
Burns, 2013	0.93	0.38	cbt	cau	ppd	Adults	29.15	1.00	−	ind	12	UK	+	+	sr	+	+
Cooper, 2003 – cbt	0.43	0.29	cbt	cau	ppd	Adults	27.69	1.00	+	ind	10	UK	+	+	sr	−	−
Cooper, 2003 – dyn	0.54	0.29	dyn	cau	ppd	Adults	27.69	1.00	+	ind	10	UK	+	+	sr	−	−
Cooper, 2003 – sup	0.25	0.28	sup	cau	ppd	Adults	27.69	1.00	+	ind	10	UK	+	+	sr	−	−
Davoudi, 2020	0.71	0.32	3rd	cau	genmed	Adults	57.32	0.53	−	ind	8	Other	+	−	+	−	−
Dekker, 2012	0.42	0.31	cbt	cau	genmed	Elderly	66.00	0.74	−	ind	1	North A	+	+	sr	+	−
Desautels, 2017	1.07	0.38	cbt	wl	genmed	Elderly	57.10	1.00	−	ind	8	North A	+	+	+	+	−
Doering, 2013	1.04	0.28	cbt	cau	genmed	Elderly	63.62	0.31	+	ind	8	North A	+	−	sr	+	−
Evans, 1995 – cbt	0.57	0.35	cbt	other	genmed	Adults	53.92	0.35	−	grp	8	North A	−	−	sr	−	−
Evans, 1995 – sup	0.98	0.37	sup	other	genmed	Adults	53.92	0.35	−	grp	8	North A	−	−	sr	−	−
Farshi, 2020	0.49	0.31	sup	cau	genmed	Adults	34.40	1.00	−	grp	7	Other	+	−	+	+	−
Freedland, 2009 – cbt	0.89	0.28	cbt	cau	genmed	Elderly	60.65	0.50	+	ind	10	North A	+	+	+	+	−
Freedland, 2009 – ssm	0.47	0.27	sup	cau	genmed	Elderly	60.65	0.50	+	ind	10	North A	+	+	+	+	−
Fuhr, 2019	0.21	0.13	cbt	cau	ppd	Adults	25.18	1.00	−	ind	10	Other	−	+	+	+	+
Furukawa, 2012	0.67	0.19	cbt	wl	other	Adults	39.35	0.22	−	tel	7	East A	+	+	sr	+	+
Gellis, 2008	1.94	0.31	pst	cau	other	Elderly	77.40	0.87	−	ind	6	North A	+	−	+	−	−
Goodman, 2015	0.03	0.30	other	cau	ppd	Adults	30.69	1.00	−	ind	8	North A	−	+	sr	+	−
Herrmann-Lingen, 2016	0.00	0.08	dyn	cau	genmed	Elderly	59.20	0.21	−	other	16	EU	+	−	+	+	+
Holden, 1989	0.72	0.32	sup	cau	ppd	Adults	26.00	1.00	+	ind	8	UK	+	−	+	−	−
Hou, 2014	0.71	0.14	cbt	cau	ppd	Adults	28.00	1.00	+	other	19	East A	−	−	sr	−	−
Joling, 2011	0.10	0.15	cbt	cau	gp	Elderly	81.45	0.74	−	gsh	3	EU	+	+	sr	+	−
Kamga, 2017	0.43	0.24	cbt	wl	genmed	Elderly	76.00	0.63	−	gsh	8	North A	+	+	sr	+	+
Kim, 2018	0.82	0.27	cbt	cau	genmed	Adults	48.00	1.00	−	other	7	East A	+	−	sr	−	−
Lamers, 2010	0.26	0.13	cbt	cau	genmed	Elderly	70.70	0.47	+	ind	4	EU	+	+	sr	+	−
Lerner, 2015	0.60	0.10	cbt	cau	other	Adults	54.70	0.72	+	tel	8	North A	+	−	sr	+	−
Lexis, 2011	0.28	0.17	pst	cau	other	Adults	47.74	0.39	−	ind	8	EU	+	−	sr	+	−
Lloyd-Williams, 2018	0.13	0.34	other	cau	genmed	Elderly	65.10	0.71	−	ind	1	UK	−	+	sr	+	−
Lund, 2019	0.19	0.11	other	cau	ppd	Adults	27.00	1.00	−	ind	6	Other	+	+	+	+	+
Lundgren, 2016	0.27	0.28	cbt	other	genmed	Elderly	62.90	0.42	−	gsh	7	EU	+	+	sr	+	+

Study	g	s.e.	therapy	Ctr	settings	age group	M age	Pr wom	Dx	Frm	Nsess	country	SG	AC	BA	IOD	SOR
Lynch, 1997	0.90	0.42	pst	cau	gp	Adults	48.30	0.86	−	tel	6	North A	−	−	sr	−	−
Lynch, 2004 – cau	0.02	0.42	pst	cau	gp	Adults	38.50	0.83	−	tel	6	North A	−	−	+	−	−
Milgrom, 2005 – cbt	0.48	0.51	cbt	cau	ppd	Adults	29.70	1.00	+	other	9	Aus	+	+	sr	+	−
Milgrom, 2005 – grp sup	0.70	0.51	sup	cau	ppd	Adults	29.70	1.00	+	other	9	Aus	+	+	sr	+	−
Milgrom, 2005 – ind sup	0.55	0.50	sup	cau	ppd	Adults	29.70	1.00	+	other	9	Aus	+	+	sr	+	−
Milgrom, 2011 – nurse	0.72	0.36	cbt	cau	ppd	Adults	31.48	1.00	−	ind	5	Aus	+	+	sr	+	−
Milgrom, 2011 – psy	0.06	0.35	cbt	cau	ppd	Adults	31.48	1.00	−	ind	5	Aus	+	+	sr	+	−
Mossey, 1996	0.21	0.27	ipt	cau	genmed	Elderly	71.00	0.78	−	ind	10	North A	−	−	sr	−	−
Neugebauer, 2006	0.15	0.44	ipt	cau	other	Adults	29.70	1.00	−	tel	6	North A	−	−	+	+	−
Ngai, 2015	0.42	0.07	cbt	cau	ppd	Adults	30.80	1.00	−	tel	5	East A	+	+	sr	+	−
Nollett, 2016 – cau	−0.07	0.31	pst	cau	genmed	Elderly	70.26	0.59	−	ind	6	UK	+	+	sr	−	−
O'Mahen, 2013 – I	3.66	0.44	cbt	cau	ppd	Adults	27.05	1.00	+	ind	8	UK	+	+	sr	+	−
O'Neil, 2014	0.12	0.18	cbt	cau	genmed	Elderly	59.96	0.25	−	tel	8	Aus	+	+	+	+	−
Onuigbo, 2019	1.74	0.29	cbt	cau	other	Adults	25.43	0.54	−	grp	14	Other	+	+	sr	+	−
Pibernik, 2015	0.08	0.18	cbt	cau	genmed	Elderly	58.12	0.54	−	grp	6	EU	+	+	sr	+	−
Richards, 2018	−0.42	0.38	bat	cau	genmed	Elderly	65.30	0.48	−	ind	8	UK	+	+	sr	+	+
Salamanca-Sanabria, 2020	0.90	0.26	cbt	wl	other	Adults	22.15	0.72	−	gsh	7	Other	+	−	+	−	−
Schulberg, 1996	0.44	0.15	ipt	cau	gp	Adults	38.09	0.83	+	ind	16	North A	−	−	+	+	−
Sheeber, 2012	0.84	0.25	cbt	wl	other	Adults	31.00	1.00	−	gsh	6	North A	+	+	sr	+	−
Sikander, 2019	0.27	0.10	cbt	cau	ppd	Adults	27.00	1.00	−	other	11	Other	+	+	+	+	+
Strong, 2008	0.37	0.14	pst	cau	genmed	Elderly	56.60	0.71	+	ind	10	UK	+	+	sr	+	−
Van Schaik, 2006	0.07	0.17	ipt	cau	gp	Elderly	67.93	0.69	+	ind	8	EU	+	−	−	+	−
Vázquez, 2020 – bat	1.25	0.21	cbt	cau	other	Adults	54.00	0.91	−	other	5	EU	+	+	+	+	+
Vázquez, 2020 – cbt	1.14	0.21	bat	cau	other	Adults	54.00	0.91	−	other	5	EU	+	+	+	+	+
Verduyn, 2003 – cbt	0.20	0.42	cbt	cau	other	Adults	29.81	1.00	−	grp	16	UK	−	+	+	−	−
Verduyn, 2003 – sup	0.14	0.43	sup	cau	other	Adults	29.81	1.00	−	grp	16	UK	−	+	+	−	−
Wickberg, 1996	1.33	0.47	sup	cau	ppd	Adults	28.40	1.00	+	ind	6	EU	−	−	+	−	−
Xie, 2019	0.68	0.24	bat	cau	other	Elderly	71.90	0.59	−	grp	8	East A	+	−	sr	−	−
Zhao, 2019	0.43	0.11	other	cau	genmed	Adults	30.52	1.00	−	grp	4	East A	+	−	sr	−	−

Note. Indicators of the columns: *g*, effect size; s.ε., standard error of effect size; therapy, type of therapy (cbt, cognitive behavior therapy; bat, behavioral activation therapy; 3rd, 3rd wave cognitive behavioral therapy; pst, problem-solving therapy; dyn, psychodynamic therapy; ipt, interpersonal psychotherapy; sup, non-directive supportive therapy; other, other type of therapy); Ctr, control group (wl, waiting list; cau, care-as-usual; other, other control condition); settings, setting category (ppd, perinatal care; genmed, general medical care; gp, general practitioner care (primary care); other, in another setting); age group, adults or elderly; M age, mean age; Pr wom, proportion of women; Dx, diagnostic interview (+, yes; −, no); Frm, format (ind, individual; grp, group; gsh, guided self-help; tel, telephone; other, other/mixed format); Nsess, number of sessions; country, country of trial (North A, North America; Aus, Australia; East A, East Asia; EU, Europe; UK, United Kingdom; other, other countries); SG, sequence generation [positive or negative (negative includes unclear)]; AC, allocation concealment [positive or negative (negative includes unclear)]; BA, blinded assessment [positive or negative (negative includes unclear); sr, self-report]; IOD, incomplete outcome data [positive or negative (negative includes unclear)]; SOR, selective outcome reporting [positive or negative (negative includes unclear)].
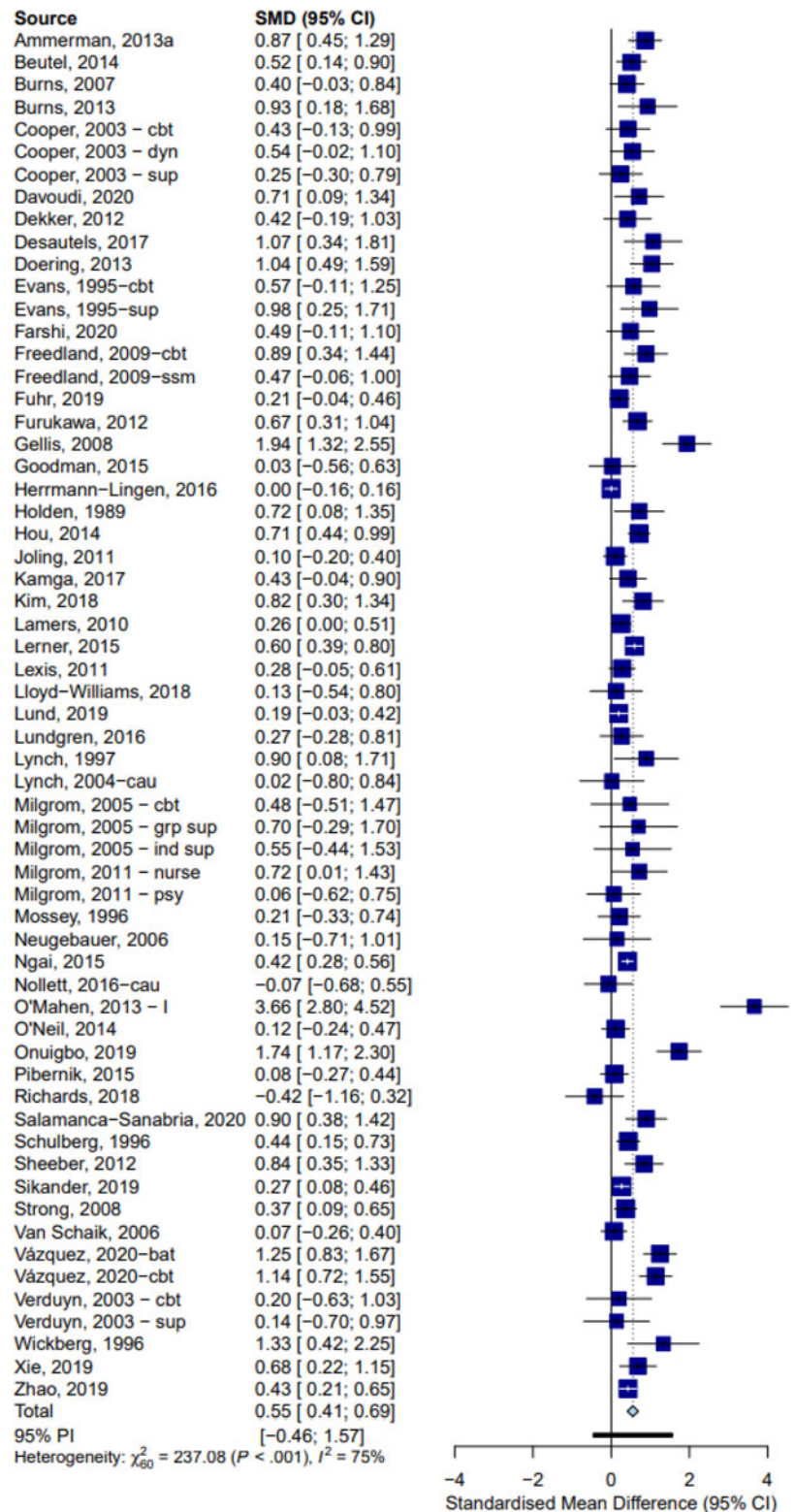
Source	SMD (95% CI)
Ammerman, 2013a	0.87 [0.45; 1.29]
Beutel, 2014	0.52 [0.14; 0.90]
Burns, 2007	0.40 [−0.03; 0.84]
Burns, 2013	0.93 [0.18; 1.68]
Cooper, 2003 − cbt	0.43 [−0.13; 0.99]
Cooper, 2003 − dyn	0.54 [−0.02; 1.10]
Cooper, 2003 − sup	0.25 [−0.30; 0.79]
Davoudi, 2020	0.71 [0.09; 1.34]
Dekker, 2012	0.42 [−0.19; 1.03]
Desautels, 2017	1.07 [0.34; 1.81]
Doering, 2013	1.04 [0.49; 1.59]
Evans, 1995−cbt	0.57 [−0.11; 1.25]
Evans, 1995−sup	0.98 [0.25; 1.71]
Farshi, 2020	0.49 [−0.11; 1.10]
Freedland, 2009−cbt	0.89 [0.34; 1.44]
Freedland, 2009−ssm	0.47 [−0.06; 1.00]
Fuhr, 2019	0.21 [−0.04; 0.46]
Furukawa, 2012	0.67 [0.31; 1.04]
Gellis, 2008	1.94 [1.32; 2.55]
Goodman, 2015	0.03 [−0.56; 0.63]
Herrmann−Lingen, 2016	0.00 [−0.16; 0.16]
Holden, 1989	0.72 [0.08; 1.35]
Hou, 2014	0.71 [0.44; 0.99]
Joling, 2011	0.10 [−0.20; 0.40]
Kamga, 2017	0.43 [−0.04; 0.90]
Kim, 2018	0.82 [0.30; 1.34]
Lamers, 2010	0.26 [0.00; 0.51]
Lerner, 2015	0.60 [0.39; 0.80]
Lexis, 2011	0.28 [−0.05; 0.61]
Lloyd−Williams, 2018	0.13 [−0.54; 0.80]
Lund, 2019	0.19 [−0.03; 0.42]
Lundgren, 2016	0.27 [−0.28; 0.81]
Lynch, 1997	0.90 [0.08; 1.71]
Lynch, 2004−cau	0.02 [−0.80; 0.84]
Milgrom, 2005 − cbt	0.48 [−0.51; 1.47]
Milgrom, 2005 − grp sup	0.70 [−0.29; 1.70]
Milgrom, 2005 − ind sup	0.55 [−0.44; 1.53]
Milgrom, 2011 − nurse	0.72 [0.01; 1.43]
Milgrom, 2011 − psy	0.06 [−0.62; 0.75]
Mossey, 1996	0.21 [−0.33; 0.74]
Neugebauer, 2006	0.15 [−0.71; 1.01]
Ngai, 2015	0.42 [0.28; 0.56]
Nollett, 2016−cau	−0.07 [−0.68; 0.55]
O'Mahen, 2013 − I	3.66 [2.80; 4.52]
O'Neil, 2014	0.12 [−0.24; 0.47]
Onuigbo, 2019	1.74 [1.17; 2.30]
Pibernik, 2015	0.08 [−0.27; 0.44]
Richards, 2018	−0.42 [−1.16; 0.32]
Salamanca−Sanabria, 2020	0.90 [0.38; 1.42]
Schulberg, 1996	0.44 [0.15; 0.73]
Sheeber, 2012	0.84 [0.35; 1.33]
Sikander, 2019	0.27 [0.08; 0.46]
Strong, 2008	0.37 [0.09; 0.65]
Van Schaik, 2006	0.07 [−0.26; 0.40]
Vázquez, 2020−bat	1.25 [0.83; 1.67]
Vázquez, 2020−cbt	1.14 [0.72; 1.55]
Verduyn, 2003 − cbt	0.20 [−0.63; 1.03]
Verduyn, 2003 − sup	0.14 [−0.70; 0.97]
Wickberg, 1996	1.33 [0.42; 2.25]
Xie, 2019	0.68 [0.22; 1.15]
Zhao, 2019	0.43 [0.21; 0.65]
Total	0.55 [0.41; 0.69]
95% PI	[−0.46; 1.57]

Heterogeneity: $\chi^2_{60} = 237.08$ ($P < .001$), $I^2 = 75\%$

Fig. 2. Forest plot of psychotherapy *v.* control conditions for adult depression: Hedges' *g*.

Standardised Mean Difference (95% CI)

analyses to examine the effect size of these studies. We found this did not change the overall pooled effect size and the level of heterogeneity (Table 2).

We found strong indications for publication bias. The funnel plot is presented in online Supplementary Appendix F.1. Egger's test of the intercept was significant (intercept: 1.53; 95% CI 0.55–2.51; *p* = 0.003), and Duval and Tweedie's trim-and-fill procedure resulted in 14 imputed studies. Adjusted effect size decreased (*g* = 0.35; 95% CI 0.17–0.52; NNT = 9.29; *p* < 0.001).

Subgroup and meta-regression analyses

In subgroup analyses (Table 2), different settings significantly moderated effect sizes (*p* = 0.017). The studies in which patients

Table 2. Effects of psychotherapies compared with control groups: Hedges' g[a]

	N (comparisons)	Effect size		Heterogeneity			
		g	95% CI	I^2	95% CI	p[b]	NNT
All comparisons	61	0.55	0.41–0.69	75	68–80		5.46
One ES per study							
Largest	52	0.57	0.41–0.73	77	70–82		5.29
Smallest	52	0.53	0.37–0.69	76	69–82		5.71
Adjusted for publication bias (add 14)	75	0.35	0.17–0.52	83	79–86		9.29
Removing nine outliers	52	0.47	0.39–0.55	30	2–51		6.64
Only low ROB	31	0.57	0.33–0.82	82	75–87		5.25
Subgroup analyses							
Setting						0.017	
General medical	24	0.40	0.26–0.53	54	27–71		8.01
Primary care	5	0.26	−0.004 to 0.52	37	0–77		12.94
Perinatal	19	0.65	0.31–0.98	78	67–86		4.56
Other	13	0.83	0.53–1.14	76	60–86		3.43
Age group						0.050	
Adults	40	0.65	0.48–0.83	73	62–80		4.54
Elderly	21	0.38	0.18–0.59	71	55–81		8.35
Diagnosis						0.209	
Cut-off	41	0.49	0.34–0.63	73	63–80		6.31
Mood disorder	20	0.70	0.40–1.01	76	63–84		4.15
Risk of bias						0.844	
Low	31	0.57	0.34–0.81	82	75–87		5.25
High	30	0.55	0.40–0.69	53	29–69		5.56

ES, effect size; NNT, numbers-needed-to-treat; ROB, risk of bias.
[a]According to the random effects model.
[b]The p value indicates whether the difference between subgroups is significant.

were recruited in primary care resulted in a lowest effect size ($g = 0.26$) compared with studies conducted in other settings (perinatal care: $g = 0.65$; general medical care: $g = 0.40$; non-specific settings: $g = 0.83$). We further conducted a post-hoc analysis to assess whether the number of sessions, which is related to the intensity of the psychotherapy, is a significant moderator for the effect sizes of psychotherapy. We categorized the number of sessions into three categories (6 or fewer; 7–11; 12 or more). But the result was not significant ($p = 0.61$) (online Supplementary Appendix K). Besides, we found no evidence indicating that effect sizes differed significantly across age group (adults *v.* elderly, $p = 0.050$), diagnosis at baseline (cut-off score *v.* diagnostic interview, $p = 0.209$), or risk of bias (low *v.* high, $p = 0.844$).

Additionally, we conducted multivariate meta-regression analyses. Because of the relatively small number of studies, only one available continuous variable (publication year) and five categorical variables [setting (general medical care *v.* other), age group (older *v.* younger adults), diagnosis at baseline (clinical interview *v.* score above a cut-off on a self-report scale), risk of bias (low *v.* high), and type of psychotherapy (cognitive behavior therapy *v.* other)] were entered as predictors. As can be seen in

online Supplementary Appendix G, no significant predictors were found (p values >0.19).

Study dropout and long-term effects

We found no significant difference between treatment and control conditions for study dropout (RR = 1.16; 95% CI 1.00–1.35; I^2 = 37%; 95% CI 12–54; $p = 0.054$). Twenty-six studies assessed the long-term effects of psychotherapies. We classified the follow-up length into four groups: 6–8 months, 9–12 months, 13–24 months, and more than 24 months after baseline (online Supplementary Appendix H). A small-to-moderate effect size was found for studies assessing the effects of psychotherapy for 6–8 months ($g = 0.33$; 95% CI 0.14–0.52; NNT = 9.83; I^2 = 75%; 95% CI 63–83). We also found a small but significant effect size for studies assessing the longer-term outcomes for 9–12 months ($g = 0.24$; 95% CI 0.11–0.37; NNT = 13.99; I^2 = 31%; 95% CI 0–64) after baseline and more than 24 months after baseline ($g = 0.18$; 95% CI 0.15–0.21; NNT = 18.85; I^2 = 0%; 95% CI 0) with low heterogeneity. Studies examined effects at 13–24 months did not point at significant effects of the interventions ($g = 0.32$; 95% CI −0.19 to 0.82; NNT = 10.23; I^2 = 92%; 95% CI 87–96).

Proportion of participants randomized in the trial from the number of participants who completed the depression questionnaires

The pooled proportion of participants randomized in the trial from the total number of participants who completed the depression questionnaires was 0.13 (95% CI 0.08–0.19), with high heterogeneity ($I^2 = 99.4\%$; 95% CI 99.4–99.5). The forest plot is given in online Supplementary Appendix I.2. Removal of 38 outliers affected the proportion only marginally (proportion = 0.12; 95% CI 0.10–0.15). The studies with low risk of bias resulted in a similar proportion (proportion = 0.14; 95% CI 0.07–0.26), and heterogeneity was still high ($I^2 = 99.5\%$; 95% CI 99.5–99.6) (Table 3). Egger's test of the intercept was not significant (intercept: −2.84; $t = 1.61$, $p = 0.114$). According to the funnel plot (online Supplementary Appendix F.3), we found no indication of publication bias. Duval and Tweedie's trim-and-fill method resulted in seven imputed studies. Adjusted proportion decreased to 0.09 (95% CI 0.05–0.14).

In subgroup analyses (Table 3), we found several significant differences between subgroups. We found significant difference among settings ($p = 0.007$). The pooled proportion in general medical care (proportion = 0.20) was higher than that in primary care (proportion = 0.09), perinatal care (proportion = 0.06), and other settings (proportion = 0.13). The proportion was significantly lower in people with a diagnosed mood disorder (proportion = 0.04) compared to those scoring above a cut-off on a self-report depression measure (proportion = 0.18) ($p < 0.0001$). However, the proportion was not significantly associated with the age group ($p = 0.244$) and the risk of bias ($p = 0.457$).

Furthermore, we conducted the same analysis for the proportion of participants randomized in the trial from the number of participants who received the depression questionnaires. The results were reported in online Supplementary Appendix J. The forest plot is given in online Supplementary Appendix I.1. The funnel plot is given in online Supplementary Appendix F.2.

Discussion

This is the first meta-analysis that examined the psychotherapy effectiveness among patients who do not actively seek help in all settings. Based on 52 studies with 61 comparisons, we found that psychological treatments had a significant, moderate-to-large effect on reducing depressive symptoms compared to control groups ($g = 0.55$). Although significant publication bias was identified, the effects remained small but significant after adjustment for the bias ($g = 0.35$). Limiting to studies with high quality did not change the effects ($g = 0.57$). The effects were retained at 12 months' follow-up, and it was small but significant ($g = 0.24$).

An earlier systematic review examining the interventions for adults found indications that actively identifying people not seeking help in primary care resulted in a small and not significant effect size ($d = 0.13$) (Cuijpers et al., 2009). However, we could not confirm that in the current meta-analysis. The current meta-analysis expanded to different settings and found a moderate-to-high effect size. The findings of the previous meta-analysis were based on a small number of trials and the quality of these trials was suboptimal. Therefore, we have more

Table 3. Pooled proportion of participants in the trial from the number of participants who completed the screening tools

	Proportion			Heterogeneity		
	N (studies)	Prop	95% CI	I^2	95% CI	*p*
All comparisons	52	0.13	0.08–0.19	99.4	99.4–99.5	
Adjusted for publication bias (add 7)	59	0.09	0.05–0.14	99.4	99.4–99.5	
Removing 38 outliers	14	0.12	0.10–0.15	93.3	90.5–95.4	
Only low risk of bias	26	0.14	0.07–0.26	99.5	99.5–99.6	
Subgroup analyses						
Setting						*0.007*
General medical	22	0.20	0.14–0.28	99.2	99.1–99.3	–
Primary care	5	0.09	0.03–0.27	99.5	99.4–99.6	–
Perinatal	14	0.06	0.04–0.11	99.5	99.4–99.5	–
Other	11	0.13	0.06–0.27	99.6	99.5–99.7	–
Age group						0.244
Adults	32	0.11	0.08–0.16	99.5	99.4–99.5	–
Elderly	20	0.15	0.10–0.23	99.2	99.0–99.3	–
Diagnosis						*<0.0001*
Cut-off	37	0.18	0.13–0.25	99.4	99.3–99.5	–
Mood disorder	13	0.04	0.03–0.06	98.5	98.3–98.9	–
Risk of bias						0.457
Low	26	0.14	0.10–0.23	99.5	99.5–99.6	–
High	26	0.11	0.07–0.15	99.2	99.1–99.3	–

Prop, proportion.

faith in the current findings, because of the broader set of studies (not only primary care), the larger number of studies and the larger number of studies with high quality. We did find that the setting where the recruitment was conducted is a significant moderator. It suggests that the effects of psychotherapy could differ considerably across settings. Only five studies were conducted in primary care, and it yielded a small and non-significant effect ($g = 0.26$), which was consistent with the previous study. One plausible explanation is because of low power (only a few numbers of studies were included). It is also possible that the depressive symptoms of patients in primary care may have had less severe conditions than those in secondary care-based studies, and little improvement is possible when they get treatments (Bortolotti, Menchetti, Bellini, Montaguti, & Berardi, 2008; Schwenk, Coyne, & Fechner-Bates, 1996; Simon & VonKorff, 1995). Another potential explanation could be that screening tools are less valid in primary care patients and may not be sensitive enough to detect the change between pre- and post-tests (Cuijpers et al., 2009; Thombs et al., 2011). But we checked the instruments used in five primary care studies, and we found that two studies used the Center for Epidemiologic Studies Depression Scale (CES-D), one study used the Primary Care Evaluation of Mental Disorders Patient Questionnaire (PRIME-MD PQ), one study used the medical outcomes study depression screening inventory, and one study used Geriatric Depression Scale. All of these screening instruments are validated, but only one study used Patient Health Questionnaire (PHQ) specifically designed to screen depression in primary care. Besides, our findings could also be explained by the intensity of the psychotherapy (Cape, Whittington, Buszewicz, Wallace, & Underwood, 2010). A previous study found that psychotherapy provided within primary care settings for depression is usually brief (Thomas & Corney, 1992). A common treatment length in the UK is six sessions (Stiles, Barkham, Connell, & Mellor-Clark, 2008). Following this, we conducted a post-hoc analysis for the number of sessions, which is related to the intensity of the psychotherapy, but we found no evidence indicating that effect sizes differed significantly across the number of sessions ($p = 0.61$). But we did find the length of the study may be a factor that limits the potential effect sizes in primary care. In the current meta-analysis with five studies conducted in primary care, three studies had six or fewer sessions, two studies had more than seven sessions. Additionally, the technical competence of the therapists or training of therapists may be related to the small effect size in primary care. However, we have not very extensively explored what kind of clinicians delivered the therapies because most trials have a mix of professionals and paraprofessionals (Cuijpers, Quero, Dowrick, & Arroll, 2019). However, the result should be interpreted with caution due to a limited number of studies. Besides, according to a previous meta-analysis, the overall prevalence of depression has been estimated to be 19.5% in primary care (Mitchell, Vaze, & Rao, 2009). However, consistent with depressed patients in other settings, this large proportion of patients is not likely to be actively seeking help. Therefore, appropriate treatments for these clients should be a public health priority. More studies are needed to explore the true effect size of psychotherapies and to draw firm conclusions on the benefits of treatment effectiveness among those not seeking help in primary care.

We also found that in different settings, the effects of psychotherapies in people not actively seeking help are consistent with studies examining the effects of psychotherapies in all kinds of participants (Cuijpers, 2017; Cuijpers, Quero, Papola, Cristea, & Karyotaki, 2021c). The psychotherapy for people not actively seeking help in perinatal care resulted in a moderate-to-high effect size ($g = 0.65$) (Cuijpers, Brännmark, & van Straten, 2008), and in general medical care yielded in a small-to-moderate effect size ($g = 0.40$) (van Straten, Geraedts, Verdonck-de Leeuw, Andersson, & Cuijpers, 2010). This implies that in most settings with exception of primary care, psychological treatments are effective for people with depression who do not actively seek help, and the effects are sustained in the 12 months' follow-up. It is widely recognized that depression accompanying medical disorders is associated with lower compliance to treatment, delayed recovery, and increased mortality (Baune, Adrian, & Jacobi, 2007; Evans & Charney, 2003). As for women during pregnancy who are depressed, depression may lead to serious health consequences for both the mother and her offspring (Dagher, Bruckheim, Colpe, Edwards, & White, 2021; Field, 2017). Therefore, treating comorbid depression should be one of the priorities in these settings, and our findings are good news for depressive patients who do not actively seek help in these settings. It is possible that if they are actively identified and provided further treatment, they could benefit from it.

Our secondary goal was to explore the proportions of patients who completed the screening questionnaire who end up in the clinical trial. We found that 13% of patients who completed the screening questionnaire would meet the criteria for depression and agree to be randomized in the trial. Additionally, the setting is a significant moderator. The proportion of patients randomized from screening in general medical care was 20%, the proportion in primary care was 9%, and in perinatal care 6%. This is not surprising because the prevalence of depression is higher for patients with medical disorders (Katon, 1996; van Straten et al., 2010).

However, the results of the current study should be interpreted with caution considering several limitations. One important limitation is that we demonstrated strong indications for publication bias in the present study. After adjustment for this bias, the effect found for the psychotherapies was decreased ($g = 0.35$), although it remained statistically significant. It indicates that the actual effects might be smaller than the primary analyses indicate. Second, there was a significant and high level of heterogeneity among the studies. In the primary analyses, we found that nine outliers contributed to heterogeneity. However, the level of heterogeneity was extremely high in the analyses of secondary outcomes (proportions), which is commonly found rather than an exception in the meta-analyses of proportions (Cuijpers et al., 2021a; Lim et al., 2018; Sheldon et al., 2021). A possible explanation is that the included studies varied widely in terms of the recruitment for people not actively seeking help and numbers of enrolled, screened, and randomized. Therefore, the resulting proportions differ so much across studies. The high level of heterogeneity implies that it cannot be reliably predicted what the true mean proportion is. Third, it is important to note that the quality of half of the included studies was not optimal. We tried to minimize the impact of study quality by limiting the analyses to studies with a low risk of bias, but the effect size was still moderate-to-high ($g = 0.57$). Although the outcomes were comparable, this still means that the results should be interpreted with caution. Even if it might be challenging to examine, the quality in which psychotherapy is delivered is an important factor that should also be considered, because it might increase the chances of inflated treatment effects in comparison with the true effects in the population (Cuijpers et al., 2020; Miguel et al., 2021).

However, since the subgroup analyses (grouped by the risk of bias assessment) did not show significant differences in effect size ($p = 0.844$), this should not be a big concern. Fourth, as mentioned before, we found few studies actively identifying people who do not seek help in primary care to the trial. Therefore, the generalizability of the results to primary care is limited. Fifth, we found that the number of sessions, which is related to the intensity of the psychotherapy, may be related to the small effect size in primary care, but intensity could be also the therapy duration in months or whether the sessions were biweekly or once a week. Unfortunately, since our database does not currently include that information, we cannot explore more about the intensity of psychotherapy.

Despite these limitations, the present study suggests that when people who do not seek help are actively identified and offered treatments, psychotherapies have the potential to decrease depressive symptoms, especially in perinatal care and general medical care settings. Our results of a moderate-to-high and sustained effect can encourage physicians and researchers to identify depressed patients and offer them timely treatment. However, the evidence is not conclusive considering the publication bias, high levels of heterogeneity, and the risk of bias in the half of included studies. The overall weakness of the results suggests that more high-quality trials exploring the effects of psychological treatments of depression among people not actively seeking help are urgently needed. Besides, since few included studies were conducted in primary care, our results of the effects of psychotherapy for depressed primary care patients were not stable. Therefore, further research concerning the large pool of primary care patients who have depression and might not actively seek help might be necessary.

References

Andrade, L. H., Alonso, J., Mneimneh, Z., Wells, J. E., Al-Hamzawi, A., Borges, G., … Kessler, R. C. (2014). Barriers to mental health treatment: Results from the WHO World Mental Health surveys. *Psychological Medicine*, *44*(6), 1303–1317. https://doi.org/10.1017/S0033291713001943.

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, *22*(4), 153–160.

Baune, B. T., Adrian, I., & Jacobi, F. (2007). Medical disorders affect health outcome and general functioning depending on comorbid major depression in the general population. *Journal of Psychosomatic Research*, *62*(2), 109–118. https://doi.org/10.1016/j.jpsychores.2006.09.014.

Berto, P., D'Ilario, D., Ruffo, P., Virgilio, R. D., & Rizzo, F. (2000). Depression: Cost-of-illness studies in the international literature, a review. *The Journal of Mental Health Policy and Economics*, *3*(1), 3–10. https://doi.org/10.1002/1099-176x(200003)3:1<3::aid-mhp68>3.0.co;2-h.

Bortolotti, B., Menchetti, M., Bellini, F., Montaguti, M. B., & Berardi, D. (2008). Psychological interventions for major depression in primary care: A meta-analytic review of randomized controlled trials. *General Hospital Psychiatry*, *30*(4), 293–302. https://doi.org/10.1016/j.genhosppsych.2008.04.001.

Cape, J., Whittington, C., Buszewicz, M., Wallace, P., & Underwood, L. (2010). Brief psychological therapies for anxiety and depression in primary care: Meta-analysis and meta-regression. *BMC Medicine*, *8*(1), 1–13. https://doi.org/10.1186/1741-7015-8-38.

Cuijpers, P. (2017). Four decades of outcome research on psychotherapies for adult depression: An overview of a series of meta-analyses. *Canadian Psychology*, *58*(1), 7–19. https://doi.org/10.1037/cap0000096.

Cuijpers, P. (2021). Indirect prevention and treatment of depression: An emerging paradigm? *Clinical Psychology in Europe*, *3*(4), 1–9. https://doi.org/https://doi.org/10.32872/cpe.6847.

Cuijpers, P., Brännmark, J. G., & van Straten, A. (2008). Psychological treatment of postpartum depression: A meta-analysis. *Journal of Clinical Psychology*, *64*(1), 103–118. https://doi.org/10.1002/jclp.20432.

Cuijpers, P., Karyotaki, E., Ciharova, M., Miguel, C., Noma, H., & Furukawa, T. A. (2021a). The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatrica Scandinavica*, *144*(3), 288–299. https://doi.org/10.1111/acps.13335.

Cuijpers, P., Karyotaki, E., de Wit, L., & Ebert, D. D. (2020). The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. *Psychotherapy Research*, *30*(3), 279–293. https://doi.org/10.1080/10503307.2019.1649732.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & Van Straten, A. (2014a). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders*, *159*, 118–126. https://doi.org/10.1016/j.jad.2014.02.026.

Cuijpers, P., Quero, S., Dowrick, C., & Arroll, B. (2019). Psychological treatment of depression in primary care: Recent developments. *Current Psychiatry Reports*, *21*(12), 129. https://doi.org/10.1007/s11920-019-1117-x.

Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., … Furukawa, T. (2021b). Psychotherapies for depression: A network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry: official journal of the World Psychiatric Association (WPA)*, *20*(2), 283–293. https://doi.org/10.1002/wps.20860.

Cuijpers, P., Quero, S., Papola, D., Cristea, I. A., & Karyotaki, E. (2021c). Care-as-usual control groups across different settings in randomized trials on psychotherapy for adult depression: A meta-analysis. *Psychological Medicine*, *51*(4), 634–644. https://doi.org/10.1017/S0033291719003581.

Cuijpers, P., van Straten, A., van Schaik, A., & Andersson, G. (2009). Psychological treatment of depression in primary care: A meta-analysis. *British Journal of General Practice*, *59*(559), 120–127. https://doi.org/10.3399/bjgp09X395139.

Cuijpers, P., Vogelzangs, N., Twisk, J., Kleiboer, A., Li, J., & Penninx, B. W. (2014b). Comprehensive meta-analysis of excess mortality in depression in the general community versus patients with specific illnesses. *American Journal of Psychiatry*, *171*(4), 453–462. https://doi.org/10.1176/appi.ajp.2013.13030325.

Dagher, R. K., Bruckheim, H. E., Colpe, L. J., Edwards, E., & White, D. B. (2021). Perinatal depression: Challenges and opportunities. *Journal of Women's Health*, *30*(2), 154–159. https://doi.org/10.1089/jwh.2020.8862.

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

Ebert, D. D., Franke, M., Kählke, F., Küchler, A. M., Bruffaerts, R., Mortier, P., … Baumeister, H. (2019). Increasing intentions to use mental health services among university students. Results of a pilot randomized controlled trial within the world health organization's world mental health international college student initiative. *International Journal of Methods in Psychiatric Research*, *28*(2), e1754. https://doi.org/10.1002/mpr.1754.

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.

Evans, D. L., & Charney, D. S. (2003). Mood disorders and medical illness: A major public health problem. *Biological Psychiatry*, 54, 177–180. https://doi.org/10.1016/s0006-3223(03)00639-5.

Field, T. (2017). Prenatal depression risk factors, developmental effects and interventions: A review. *Journal of Pregnancy and Child Health*, 04(01), 1–25. https://doi.org/10.4172/2376-127x.1000301.

Friedrich, M. J. (2017). Depression is the leading cause of disability around the world. *JAMA*, 317(15), 1517. https://doi.org/10.1001/jama.2017.3826.

Furukawa, T. A. (1999). From effect size into number needed to treat. *The Lancet*, 353(9165), 1680.

Greenberg, P. E., & Birnbaum, H. G. (2005). The economic burden of depression in the US: Societal and patient perspectives. *Expert Opinion on Pharmacotherapy*, 6(3), 369–376. https://doi.org/10.1517/14656566.6.3.369.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2022). *Doing meta-analysis with R: A hands-on guide*. Boca Raton: CRC Press.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., … Sterne, J. A. C. (2011). The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Clinical research ed.)*, 343, d5928. https://doi.org/10.1136/bmj.d5928.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560.

Katon, W. (1996). The impact of major depression on chronic medical illness. *General Hospital Psychiatry*, 18(4), 215–219. https://doi.org/10.1016/0163-8343(96)00065-5.

Keshavarz, H., Fitzpatrick-Lewis, D., Streiner, D. L., Maureen, R., Ali, U., Shannon, H. S., & Raina, P. (2013). Screening for depression: A systematic review and meta-analysis. *CMAJ Open*, 1(4), E159–E167. https://doi.org/10.9778/cmajo.20130030.

Lim, G. Y., Tam, W. W., Lu, Y., Ho, C. S., Zhang, M. W., & Ho, R. C. (2018). Prevalence of depression in the community from 30 countries between 1994 and 2014. *Scientific Reports*, 8(1), 2861. https://doi.org/10.1038/s41598-018-21243-x.

Lipson, S. K., Speer, N., Brunwasser, S., Hahn, E., & Eisenberg, D. (2014). Gatekeeper training and access to mental health care at universities and colleges. *Journal of Adolescent Health*, 55(5), 612–619. https://doi.org/10.1016/j.jadohealth.2014.05.009.

Meijer, A., Roseman, M., Delisle, V. C., Milette, K., Levis, B., Syamchandra, A., … Thombs, B. D. (2013). Effects of screening for psychological distress on patient outcomes in cancer: A systematic review. *Journal of Psychosomatic Research*, 75(1), 1–17. https://doi.org/10.1016/j.jpsychores.2013.01.012.

Miguel, C., Karyotaki, E., Cuijpers, P., & Cristea, I. A. (2021). Selective outcome reporting and the effectiveness of psychotherapies for depression. *World Psychiatry*, 20(3), 444–445. https://doi.org/10.1002/wps.20900.

Mitchell, A. J., Vaze, A., & Rao, S. (2009). Clinical diagnosis of depression in primary care: A meta-analysis. *The Lancet*, 374(9690), 609–619. https://doi.org/10.1016/S0140-6736(09)60879-5.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097.

Mojtabai, R. (2009). Unmet need for treatment of major depression in the United States. *Psychiatric Services (Washington, D.C.)*, 60(3), 297–305. https://doi.org/10.1176/appi.ps.60.3.297.

NICE (2009). Depression in adults: Recognition and management. Clinical guideline. In Clinical Guideline 90. Retrieved from www.nice.org.uk/guidance/cg90.

Olkin, I., & Hedges, L. V. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press. Retrieved from https://books.google.nl/books?id=sJhQAQAACAAJ.

Pignone, M. P., Gaynes, B. N., Rushton, J. L., Burchell, C. M., Orleans, C. T., Mulrow, C. D., & Lohr, K. N. (2002). Screening for depression in adults: A summary of the evidence for the U.S. preventive services task force. *Annals of Internal Medicine*, 136, 765–776. https://doi.org/10.7326/0003-4819-136-10-200205210-00013.

Pratt, L. A., & Brody, D. J. (2014). Depression in the U.S. household population, 2009–2012. *NCHS Data Brief 2014*, (172), 1–8.

Salerno, J. P. (2016). Effectiveness of universal school-based mental health awareness programs among youth in the United States: A systematic review. *Journal of School Health*, 86(12), 922–931. https://doi.org/10.1111/josh.12461.

Schnyder, N., Panczak, R., Groth, N., & Schultze-Lutter, F. (2017). Association between mental health-related stigma and active help-seeking: Systematic review and meta-analysis. *British Journal of Psychiatry*, 210(4), 261–268. https://doi.org/10.1192/bjp.bp.116.189464.

Schwenk, T. L., Coyne, J. C., & Fechner-Bates, S. (1996). Differences between detected and undetected patients in primary care and depressed psychiatric patients. *General Hospital Psychiatry*, 18(6), 407–415. https://doi.org/10.1016/s0163-8343(96)00062-x.

Sheldon, E., Simmonds-Buckley, M., Bone, C., Mascarenhas, T., Chan, N., Wincott, M., … Barkham, M. (2021). Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis. *Journal of Affective Disorders*, 287, 282–292. https://doi.org/10.1016/j.jad.2021.03.054.

Simon, G. E., & VonKorff, M. (1995). Recognition, management, and outcomes of depression in primary care. *Archives of Family Medicine*, 4(2), 99–105. https://doi.org/10.1001/archfami.4.2.99.

Siu, A. L., Bibbins-Domingo, K., Grossman, D. C., Baumann, L. C., Davidson, K. W., Ebell, M., … Pignone, M. P. (2016). Screening for depression in adults: US preventive services task force recommendation statement. *JAMA – Journal of the American Medical Association*, 315(4), 380–387. https://doi.org/10.1001/jama.2015.18392.

Stiles, W. B., Barkham, M., Connell, J., & Mellor-Clark, J. (2008). Responsive regulation of treatment duration in routine practice in United Kingdom primary care settings: Replication in a larger sample. *Journal of Consulting and Clinical Psychology*, 76(2), 298–305. https://doi.org/10.1037/0022-006X.76.2.298.

Thomas, R. V., & Corney, R. H. (1992). A survey of links between mental health professionals and general practice in six district health authorities. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 42(362), 358–361.

Thombs, B. D., Coyne, J. C., Cuijpers, P., de Jonge, P., Gilbody, S., Ioannidis, J. P. A., … Ziegelstein, R. C. (2011). Rethinking recommendations for screening for depression in primary care. *Canadian Medical Association Journal*, 184(4), 413–418. https://doi.org/10.1503/cmaj.111035.

Thombs, B. D., Markham, S., Rice, D. B., & Ziegelstein, R. C. (2021). Does depression screening in primary care improve mental health outcomes? *BMJ*, 374, 1–8. https://doi.org/10.1136/bmj.n1661.

Thombs, B. D., Roseman, M., Coyne, J. C., de Jonge, P., Delisle, V. C., Arthurs, E., … Ziegelstein, R. C. (2013). Does evidence support the American Heart Association's recommendation to screen patients for depression in cardiovascular care? An updated systematic review. *PLoS ONE*, 8(1), e52654. https://doi.org/10.1371/journal.pone.0052654.

Thombs, B. D., Ziegelstein, R. C., Roseman, M., Kloda, L. A., & Ioannidis, J. P. A. (2014). There are no randomized controlled trials that support the United States Preventive Services Task Force guideline on screening for depression in primary care: A systematic review. *BMC Medicine, 12*, 13. https://doi.org/10.1186/1741-7015-12-13.

van Straten, A., Geraedts, A., Verdonck-de Leeuw, I., Andersson, G., & Cuijpers, P. (2010). Psychological treatment of depressive symptoms in patients with medical disorders: A meta-analysis. *Journal of Psychosomatic Research*, 69(1), 23–32. https://doi.org/10.1016/j.jpsychores.2010.01.019.

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

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., … Collaborators, P. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6.

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Geneva, Switzerland: World Health Organization.

Xu, Z., Huang, F., Kösters, M., Staiger, T., Becker, T., Thornicroft, G., & Rüsch, N. (2018). Effectiveness of interventions to promote help-seeking for mental health problems: Systematic review and meta-analysis. *Psychological Medicine*, 48(16), 2658–2667. https://doi.org/10.1017/S0033291718001265.